

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2003, 07:20:00 ; Search time 1677 Seconds  
(without alignments)  
9660.234 Million cell updates/sec

Title: US-09-684-215A-3

Perfect score: 396

Sequence: 1 acggcgcgtccgataactt.....tggccgaggagccccggcc 396

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 288711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pln.\*
- 35: em\_htg\_rod.\*
- 36: em\_htg\_mam.\*
- 37: em\_htg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	396	100.0	1068	6	AX005788 Sequence
2	396	100.0	1143	6	AX005790 Sequence
3	396	100.0	2287	6	AR303127 Sequence
4	396	100.0	11700	1	MTCIA18B
5	396	100.0	14029	1	AE006925
6	396	100.0	343050	1	EX248334 Mycobacte
7	394.4	99.6	447	6	AR169152 Sequence
8	394.4	99.6	447	6	AR182442 Sequence
9	394.4	99.6	447	6	AR194825 Sequence
10	394.4	99.6	447	6	AR233097 Sequence
11	394.4	99.6	447	6	AR235956 Sequence
12	394.4	99.6	447	6	BD006325 Compounds
13	394.4	99.6	447	6	BD006445 Compounds
14	394.4	99.6	447	6	BD069285 Compounds
15	391.2	98.8	1872	6	AR169165 Sequence
16	391.2	98.8	1872	6	AR182455 Sequence
17	391.2	98.8	1872	6	AR194838 Sequence
18	391.2	98.8	1872	6	AR233110 Sequence
19	391.2	98.8	1872	6	AX429609 Sequence
20	391.2	98.8	1872	6	BD006338 Compounds
21	391.2	98.8	1872	6	BD006458 Compounds
22	391.2	98.8	1872	6	BD069298 Compounds
23	354	89.4	675	6	AR261272 Sequence
24	354	89.4	675	6	AX201049 Sequence
25	354	89.4	675	6	AX267848 Sequence
26	354	89.4	822	6	AR277645 Sequence
27	354	89.4	822	6	AX369152 Sequence
28	354	89.4	894	6	AX351489 Sequence
29	354	89.4	900	6	AR220690 Sequence
30	354	89.4	900	6	AX365960 Sequence
31	354	89.4	915	6	AR261281 Sequence
32	354	89.4	915	6	AX201061 Sequence
33	354	89.4	915	6	AX267860 Sequence
34	354	89.4	945	6	AR277644 Sequence
35	354	89.4	945	6	AX369151 Sequence
36	354	89.4	1012	6	AR220689 Sequence
37	354	89.4	1012	6	AX365958 Sequence
38	354	89.4	1155	6	AX369165 Sequence
39	354	89.4	1203	6	AX201078 Sequence
40	354	89.4	1203	6	AX267877 Sequence
41	354	89.4	1464	6	AR229410 Sequence
42	354	89.4	1464	6	AX156105 Sequence
43	354	89.4	1464	6	AX361955 Sequence
44	354	89.4	1557	6	AR229398 Sequence
45	354	89.4	1557	6	AX156089 Sequence

ALIGNMENTS

RESULT 1	AX005788	Sequence 907 from Patent WO9909186.	1068 bp	DNA	linear	PAT 24-AUG-2000
AX005788	LOCUS	AX005788				
DEFINITION	Sequence 907 from Patent WO9909186.					
ACCESSION	AX005788					
VERSION	AX005788.1	GI:9928795				
KEYWORDS						
SOURCE	Mycobacterium tuberculosis					
ORGANISM	Mycobacterium tuberculosis					
	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;					
	Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium					
	tuberculosis complex.					
REFERENCE	1					
AUTHORS	Portnoi, D. and Guigueno, A.					
TITLE	Polypeptide nucleic sequences exported from mycobacteria, vectors					

comprising same and uses for diagnosing and preventing tuberculosis  
 Patent: WO 9909186-A 907 25-FEB-1999;  
 PORTNOI DENIS (FR); GUIGUENO AGNES (FR)

JOURNAL  
 FEATURES source  
 Location/Qualifiers

1. .1068  
 /organism="Mycobacterium tuberculosis"

/mol\_type="genomic DNA"

/db\_xref="taxon:1773"

1. .1068

/note="50D"

/codon\_start=1

/transl\_table=11

/protein\_id="CAC05178.1"

/db\_xref="GI:9928796"

/translation="MSNSRRRLRWSSLLVLAAYGLGLATAPAAAPALSDREFAD  
 FPALPLDPSAMVAQPGVNTNKLGYNNAGCTGVIDPNGVLFNNHVIAGATD  
 INAFVSGSGQTVGVYDRTQDVAVLQLRAGGLPSAAICGGVAGPEPVVAMNSG  
 GGGTPRAVGRVALGQTVQASDLSLTGAETLNLGIQDAIOPDGGGPPVNLGQ  
 VYGMATAASDNFQSGGGGFAIPIGQAMATAGTIRSGGSPVHIGPTAFGLGVVD  
 NNGGARVQRVVSAPAAASLGISTGDIVTAVDGPINATAMADALNGHHPGDVISVT  
 WQTKSGGTGTGNVTLAEGPPA"

BASE COUNT 169 a 352 c 375 g 172 t

ORIGIN

Query Match 100.0%; Score 396; DB 6; Length 1068;

Best Local Similarity 100.0%; Pred. No. 4.2e-57;

Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGGCGGCTCCGATTAACCTCCAGCTGTCCAGGCTGGGAGGATTCGCCATTCCGATC 60  
 DB 670 ACGGCGGCTCCGATTAACCTCCAGCTGTCCAGGCTGGGAGGATTCGCCATTCCGATC 729  
 QY 61 GGGCAGGCGATGGCGATCGCGGCCAGATCCGATCGGTTGGGGGTTCACCCACCGTTTAT 120  
 DB 730 GGGCAGGCGATGGCGATCGCGGCCAGATCCGATCGGTTGGGGGTTCACCCACCGTTTAT 789  
 QY 121 ATCGGCGCTACCGCTTCTCGGCTTGGGTGTGTCGACAAACAGGCAACGGCGACGCA 180  
 DB 790 ATCGGCGCTACCGCTTCTCGGCTTGGGTGTGTCGACAAACAGGCAACGGCGACGCA 849  
 QY 181 GTCCAAACGGTGTGGGAGCGCTCGGGGCAAGTCTCGGATCTCCACCGCGGACGCTG 240  
 DB 850 GTCCAAACGGTGTGGGAGCGCTCGGGGCAAGTCTCGGATCTCCACCGCGGACGCTG 909  
 QY 241 ATCACCGGCTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGACGCGTTAAC 300  
 DB 910 ATCACCGGCTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGACGCGTTAAC 969  
 QY 301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAACCAAGTCGGGGCGGACGCGT 360  
 DB 970 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAACCAAGTCGGGGCGGACGCGT 1029  
 QY 361 ACAGGGAAGTGACATTTGGCCGAGGACCGCCCGGCC 396  
 DB 1030 ACAGGGAAGTGACATTTGGCCGAGGACCGCCCGGCC 1065

RESULT 2

AX005790

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

1

comprising same and uses for diagnosing and preventing tuberculosis

Portnoi, D. and Guigueno, A.

Polypeptide nucleic sequences exported from mycobacteria, vectors

comprising same and uses for diagnosing and preventing tuberculosis

JOURNAL

FEATURES

source

Location/Qualifiers

1. .1143

/organism="Mycobacterium tuberculosis"

/mol\_type="genomic DNA"

/db\_xref="taxon:1773"

4. .1143

/note="SEQ ID NO 50F"

/codon\_start=1

/transl\_table=11

/protein\_id="CAC05179.1"

/db\_xref="GI:9928798"

/translation="KNPAARTLKAGTISDYPDTGCGMSNRRLRWSSLLSVLAA

VGLGLATAPAAAPALSDREFADFPALPLDPSAMVAQPGVNTNKLGYNNAGCTG

VIDPNGVLFNNHVIAGATDINAFVSGSGQTVGVYDRTQDVAVLQLRAGGLPSAA

ICGGVAGPEPVVAMNSG GGGTPRAVGRVALGQTVQASDLSLTGAETLNLGIQDAIOPD

GGGPPVNLGQ VYGMATAASDNFQSGGGGFAIPIGQAMATAGTIRSGGSPVHIGPTAF

GLGVVD NNGGARVQRVVSAPAAASLGISTGDIVTAVDGPINATAMADALNGHHPGD

VISVT WQTKSGGTGTGNVTLAEGPPA"

BASE COUNT 189 a 373 c 395 g 186 t

ORIGIN

Query Match 100.0%; Score 396; DB 6; Length 1143;

Best Local Similarity 100.0%; Pred. No. 4.2e-57;

Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGGCGGCTCCGATTAACCTCCAGCTGTCCAGGCTGGGAGGATTCGCCATTCCGATC 60  
 DB 745 ACGGCGGCTCCGATTAACCTCCAGCTGTCCAGGCTGGGAGGATTCGCCATTCCGATC 804  
 QY 61 GGGCAGGCGATGGCGATCGCGGCCAGATCCGATCGGTTGGGGGTTCACCCACCGTTTAT 120  
 DB 805 GGGCAGGCGATGGCGATCGCGGCCAGATCCGATCGGTTGGGGGTTCACCCACCGTTTAT 864  
 QY 121 ATCGGCGCTACCGCTTCTCGGCTTGGGTGTGTCGACAAACAGGCAACGGCGACGCA 180  
 DB 865 ATCGGCGCTACCGCTTCTCGGCTTGGGTGTGTCGACAAACAGGCAACGGCGACGCA 924  
 QY 181 GTCCAAACGGTGTGGGAGCGCTCGGGGCAAGTCTCGGATCTCCACCGCGGACGCTG 240  
 DB 925 GTCCAAACGGTGTGGGAGCGCTCGGGGCAAGTCTCGGATCTCCACCGCGGACGCTG 984  
 QY 241 ATCACCGGCTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGACGCGTTAAC 300  
 DB 985 ATCACCGGCTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGACGCGTTAAC 1044  
 QY 301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAACCAAGTCGGGGCGGACGCGT 360  
 DB 1045 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAACCAAGTCGGGGCGGACGCGT 1104  
 QY 361 ACAGGGAAGTGACATTTGGCCGAGGACCGCCCGGCC 396  
 DB 1105 ACAGGGAAGTGACATTTGGCCGAGGACCGCCCGGCC 1140

RESULT 3

AR303127

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

1

comprising same and uses for diagnosing and preventing tuberculosis

Skeiky, F., Anderson, M. and Campos-Neto, A.

Fusion proteins of mycobacterium tuberculosis antigens and their

uses

Patent: US 6544522-A 1 08-APR-2003;

Location/Qualifiers

1. .2287

AR303127

Sequence 1 from patent US 6544522.

AR303127

GI:31691855

Unknown.

Unknown.

Unclassified.

1 (bases 1 to 2287)

Skeiky, F., Anderson, M. and Campos-Neto, A.

Fusion proteins of mycobacterium tuberculosis antigens and their

uses

Patent: US 6544522-A 1 08-APR-2003;

Location/Qualifiers

1. .2287

Source

Linear

DNA

2287 bp

PAT 12-JUN-2003

```

BASE COUNT. 372 a 717 c 814 g 381 t 3 others
ORIGIN
Query Match 100.0%; Score 396; DB 6; Length 2287;
Best Local Similarity 100.0%; Pred. No. 3.7e-57;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGGCGCGTCGCGATACCTCCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCGATC 60
DB 63 ACGGCGCGTCGCGATACCTCCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCGATC 122
QY 61 GGGCAGCGGATGCGATCGCGGCCAGATCCGATCGGGTGGGGGTCACCCACCGTTCAT 120
DB 123 GGGCAGCGGATGCGATCGCGGCCAGATCCGATCGGGTGGGGGTCACCCACCGTTCAT 182
QY 121 ATCGGGCTACCGCTTCCTCGGCTTGGGTGGTGTGTCACAAACGCGGCGGACGA 180
DB 183 ATCGGGCTACCGCTTCCTCGGCTTGGGTGGTGTGTCACAAACGCGGCGGACGA 242
QY 181 GTCCACGCGTGTGCGGAGCGCTCGCGCGGCAAGTCTCGGCATCTCCACCGCGGCGGTG 240
DB 243 GTCCACGCGTGTGCGGAGCGCTCGCGCGGCAAGTCTCGGCATCTCCACCGCGGCGGTG 302
QY 241 ATCACCGCGTGCAGCGCGCTCGCATCACTCGGCCACCGCGATGCGGCGGCGCTTAAC 300
DB 303 ATCACCGCGTGCAGCGCGCTCGCATCACTCGGCCACCGCGATGCGGCGGCGCTTAAC 362
QY 301 GGGCATCATCCGGTGCAGTCACTCTCGGTGACCTGGCAACCAAGTCGGCGGCGCGGT 360
DB 363 GGGCATCATCCGGTGCAGTCACTCTCGGTGACCTGGCAACCAAGTCGGCGGCGCGGT 422
QY 361 ACAGGGACGTCAGATTGGCGGAGGACCCCGGCC 396
DB 423 ACAGGGACGTCAGATTGGCGGAGGACCCCGGCC 458

RESULT 4
MTCI418B 11700 bp DNA linear BCT 02-SEP-2002
LOCUS Mycobacterium tuberculosis H37Rv complete genome; segment 7/162.
ACCESSION Z96071 ALI23456
VERSION Z96071.1 GI:3242254
KEYWORDS
SOURCE Mycobacterium tuberculosis H37Rv
ORGANISM Mycobacterium tuberculosis H37Rv
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
REFERENCE 1
AUTHORS Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C.,
Harris, D., Gordon, S.V., Eiglmeier, K., Gas, S., Barry III, C.E.,
Tekala, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T.,
Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S.,
Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J.,
Moule, S., Murphy, L., Oliver, S., Osborne, J., Quail, M.A.,
Rajandream, M.A., Rogers, J., Rutter, S., Seeger, K., Skelton, S.,
Squares, S., Squares, R., Sultston, J.E., Taylor, K., Whitehead, S. and
Barrell, B.G.
Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence
NATURE 393 (6685), 537-544 (1998)
JOURNAL Nature 393 (6685), 537-544 (1998)
MEDLINE 98295987
PUBMED 9634230
REFERENCE 2 (bases 1 to 11700)
AUTHORS Parkhill, J.
DIRECT SUBMISSION
Submitted (11-JUN-1998) Submitted on behalf of the Mycobacterium
tuberculosis sequencing and mapping teams, Sanger Centre, Wellcome
Trust Genome Campus, Hinxton, Cambridge CB10 1SA Unite de Genetique
Moleculaire Bacterienne, Institut Pasteur, 28 rue du Docteur Roux,
75724 Paris Cedex 15, France E-mail: parkhill@sanger.ac.uk
On Jun 20, 1998 this sequence version replaced gi:2181960.

```

Notes:  
Details of M. tuberculosis sequencing at the Sanger Centre are available on the World Wide Web.  
(URL, <http://www.sanger.ac.uk/Projects/M.tuberculosis/>) CDS have been renumbered from the original cosmid submissions but the old gene designations are in brackets after the new gene numbers.  
Gene prediction was based on a Hidden Markov Model of TB genes implemented in TBparse (Krogh) supplemented with visual inspection of positional base preference in codons, especially where there is an increase in the observed/expected third position G + C.  
CAUTION: In some cases we may not have predicted the correct initiation codon. Where possible we choose an initiation codon (atg, gtg, or ttg) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.

FEATURES

source	Location/Qualifiers
	1..11700
	/organism="Mycobacterium tuberculosis H37Rv"
	/mol_type="genomic DNA"
	/strain="H37Rv"
	/db_xref="taxon:83332"
source	1309..11293
	/organism="Mycobacterium tuberculosis H37Rv"
	/mol_type="genomic DNA"
	/strain="H37Rv"
	/db_xref="taxon:83332"
	/clone="I418B"
source	10774..>11700
	/organism="Mycobacterium tuberculosis H37Rv"
	/mol_type="genomic DNA"
	/strain="H37Rv"
	/db_xref="taxon:83332"
	/clone="CI5"
misc_feature	<1..1409
	/note="fragment designated v031. Does not represent a physical clone"
RBS	83..88
	/note="possible RBS, AAGGAG, for RV0119"
gene	97..1674
	/gene="fadD7"
CDS	97..1674
	/gene="fadD7"
	/note="RV0119, (MTV031.13-MTCI418B.01), len: 525.fadD7, PossibleCoenzyme A-ligase similar to 4-coumarate:CoA ligase of many organisms e.g. PTU39405_1 (537 aa). Contains PS00455Putative AMP-binding domain signature. FASTA scores: gp U39405 PTU39405_1 Pinus taeda xyiem 4-coumarate:CoA (537 aa).opt:483z-score: 526.1 E(): 8.3e-22; 28.2% identity in 440 aaoverlap score is 0.896"
	/codon_start=1
	/transl_table=11
	/product="fadD7"
	/protein_id="CAB09455.1"
	/db_xref="GI:3242255"
	/db_xref="SPTREMBL:C07169"
	/translation="MASDFCPRIADLVEVAATRLPEAPALVVTADRIAISHRDRLARV DELAGOLTRSGLLPGDRVALRMGSNAEEFVALLAASRADLVVPLDLPALPITEORVRS QARAGRVLLIDAGPHDRAEPTTRWHPLVNVGSDGSPGGTSLVHLDATENPATIS TGPLGRDDAMIMFTGTTGLPKMVPWTHANIASVRAITGTRLSFROATVAVMPLY HGHLIASLLATLISAGGAVSLPARGVSAHRTWDDIKAVGATWYTAFTVTHQILLERS ATEPSGRKPAALRFIRSCSAPLTAQAALQTEFAAPVAVCAFGMTATHTVTTQIEG IDQETPVYSTGLVGRSTGAQIRIVAGSDGLPLPAGAVGEIWLRTVWVRYLGDPTIT ANFTDGLWLTGDLGSLSAAGDLSIRGKELINRGSEKISPERVEGLASHPNVMEA AVFGVPHQLYGEAVAVIVPRESAPPTREELVQFCRERLAARFAEIPASQASGLPHTA KGSDDRRAVAERFHSV"
	610..645
	/gene="fadD7"
misc_feature	/note="PS00455 Putative AMP-binding domain signature, [LIVMEY] [21][STC][STAG][ST][STEI][SG].[PASLIYW][KR], info count = 22.0"
gene	complement(1675..3819)
	/gene="fusA2"





**TITLE** Whole genome comparison of Mycobacterium tuberculosis clinical and laboratory strains

**JOURNAL REFERENCE** 2 (bases 1 to 14029)

**AUTHORS** Fleischmann, R.D., Alland, D., Eisen, J.A., Carpenter, L., White, O., Peterson, J., DeBoy, R., Dodson, R., Gwinn, M., Haft, D., Hickey, E., Kolonay, J.F., Nelson, W.C., Umayam, L.A., Ermolaeva, M., Salzberg, S.L., Delcher, A., Utterback, T., Weidman, J., Khouri, H., Gill, J., Mikula, A. and Bishai, W.

**TITLE** Direct Submission

**JOURNAL** Submitted (25-APR-2001) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA

**FEATURES** location/Qualifiers

1. 14029

    /organism="Mycobacterium tuberculosis CDC1551"

    /mol\_type="genomic DNA"

    /strain="CDC1551"

    /db\_xref="taxon:83331"

    /note="clinical strain"

    /complement(99..533)

    /gene="MT0129"

    /complement(99..533)

    /gene="MT0129"

    /note="identified by Glimmer2; putative"

    /codon\_start=1

    /transl\_table=11

    /product="hypothetical protein"

    /protein\_id="AAK44353.1"

    /db\_xref="GI:13879611"

    /translation="MGEFDEKLRFAQSPVARLATSPDGTPLHVPVVFALGARRPAEA TGAIVTAVDAKRTQRLRLANLHNPRASVLVDSYADDWTQLMMWRADGVAAIHR RDEGVMRAVLLRAKYTOYOSVPLNGPVIATIAVORWASRHA"

    /complement(643..1062)

    /gene="MT0130"

    /complement(643..1062)

    /gene="MT0130"

    /note="identified by Glimmer2; putative"

    /codon\_start=1

    /transl\_table=11

    /product="hypothetical protein"

    /protein\_id="AAK44354.1"

    /db\_xref="GI:13879612"

    /translation="MASSSILPLLRHCVSDQVTVVGFDDGLGKTARRIAALVQORA IFLNDRTVAHSGDDAVLGHAVLPVGRPDPLHASSMYSILGMCOSVNGRPPDATA LVSVRCHVQTDPTDSCGGRDRPQLCAPLDYHRHH"

    /1047..1415

    /gene="MT0131"

    /1047..1415

    /gene="MT0131"

    /note="identified by match to protein family HMM"

    /codon\_start=1

    /transl\_table=11

    /product="DNA-binding protein, CopG family"

    /protein\_id="AAK44355.1"

    /db\_xref="GI:13879613"

    /translation="MTKKPRNPADYVIGDDVEYSDVLKQEEVYVDGERITDERVEOM ASLSRLAREANLIPGKSLSGSSAHPAVQVVSVAHAKLKEIARSRKMSVSKL LRPVLDFVQRETRILPRR"

    /1682..3367

    /gene="MT0132"

    /1682..3367

    /gene="MT0132"

    /note="similar to GB:AL123456; identified by sequence similarity; putative"

    /codon\_start=1

    /transl\_table=11

    /product="PE\_PGRS family protein"

    /protein\_id="AAK44356.1"

    /db\_xref="GI:13879614"

    /translation="MFAGGAGGLRCVMSFVSVAPEIVVAAATDLAGISATSAANA AAAAPTAVLAAGADEVSAIAALFSGHAQYALSAQAAAFHOQFVOTLAGGAGAYA AEAQVFEQILLAINAPTQALLGRPLINGADGAPGTQAGGAGGILYNGNGGSGA AGQNGGAGGAGLIGHGGSGAGSGAGGHWGLWNGGVGGSGAGVAGVAG

GHGGAGGAAGLWAGAGGGGNGGADANIVSGDGLGGAGGGGWLVDGAGGGHGG  
QAGGGAGGGGGGGAGAGLWGTGGAGGGLGGGGTGGPPLPGQAGMAAGGAGG  
LIGNGAGGGGCVGAGSVGAGVXGAGGNAMLIHGAGGAGGDSFANGAGGAGGAG  
GHLFGNGSGGHHGAVTAGNTGIGGAGVGGDARLIGHGAGGAGGDRAGALVGRDGG  
PGNGGAGGOLYNGGDDGCGGQAGFANNIGTGTGAGCGNGGAPAILSGNGGAGG  
GAGGAGGAGGAGGAGGAGGAGGTGGTLOAAVSLVTLFALFAGCQPGDGTGPG"

3519..4586

/gene="MT0133"

3519..4586

/gene="MT0133"

/note="identified by match to protein family HMM"

/codon\_start=1

/transl\_table=11

/product="serine protease, putative"

/protein\_id="AAK44357.1"

/db\_xref="GI:13879615"

/translation="MSNSRRSLRWSLLSVLAAGVLGLATAPAAQAPALSDREFAD  
FPALPLDSAMVAQVQGVVNTKLTGNNAGAGTGIVDPNGVLTNNHVIAGATD  
INAFSVSGGTYGVVVDVQVQLQVLRGAGLPSAAIGGVAVGEPVAVMNGSG  
QGGTPRAVGFVVALGTVQASDSLTAETLNLQIFDAAIQPGSGGPPVAVNGSQ  
VVGMTAASDNFOLSGGQGAIPIGQAMATAGQIRSGGSPTHIGPTAFGLGVVD  
NNGGARVORVVGSAAPASLIGISTGDIVTAVDGAIPINSATAMADALNCHHHPGDVISVT  
WOKSGGTRTGNVTLAEGPPA"

4695..6500

/gene="MT0134"

4695..6500

/gene="MT0134"

/note="similar to GP:2808807; identified by sequence similarity; putative"

/codon\_start=1

/transl\_table=11

/product="alpha-amylase family protein"

/protein\_id="AAK44358.1"

/db\_xref="GI:13879616"

/translation="MNEAHSVEHPVQGSHEGVVEHPDAKDFGSAALPADPTWF  
KHAVFVLYRAFFDASADSGDLRLDLRLDYLQNLGIDCIWLPPYDSELRGGYD  
IRDFYKVLPEFGTVDDFVALVDAARRGIRITLDMNTHSSEHPWFQESRRDPDGPY  
GDYVWSDTSERYTDARIIFVDTEESNMSDPVRRQYWHRRFFSHQPDLDNPAVQE  
AMIDVIRFWLGLDIDGFLDAPVYLFEREGTNCNLPETHAFKRVKRVDDPEPGVY  
LLAEANQWPGDVVEYFGDPNTGDECHMAHFPLMPRIFMAVRRESRPISEILAQTP  
PIPDMAQGIPLRNHDELTEMVTDDEERDYMAYAKDPRMKANVGIRRLAPLLDND  
RNQIETALLLSLPGSPVLYYDEIGMDVWLGDRDGVIRPQWTFDRNAGSTAN  
PRLYLPPSQDPVYQVQANVQADRTSLNLTMTLAVRRHAPAFVAGFQELGG  
SNPSVLAVPQVAGDGGTVLCVNNLSRFQPIELDLQOMNTYTPVELTGHVEFPRIG  
QVYPLTLPGHGFYWFOLTTHEVGAPPTCGGERL"

6603..7970

/gene="MT0135"

6603..7970

/gene="MT0135"

/note="similar to GP:2808801; identified by sequence similarity; putative"

/codon\_start=1

/transl\_table=11

/product="pep2 protein"

/protein\_id="AAK44359.1"

/db\_xref="GI:13879617"

/translation="MTRSDTLATKLPSDWMLPRQWYAGNRRELATVKPGYVVALRHN  
LDLVLVDTVTGATERYQVLWGDFFPASEYGTAAIGVADRTGTDALYDVAGPQF  
LLSLIVSSAVCGTSTGEVTFTRPDELPPAAOPRVCDAEQSTSVIFDRRAILKVFY  
RVSSGINPDIELNVLTRAGNPHVRLILGAYQFCGRPNRSPDIALAYLGVTEYANA  
AEGWAMATASVRLDFABGDLTAHEVGGDFAGESRGLSEAVASVATLADSLGTAAATF  
PVDRLARLSSTVAVVPELREYAPTIEQFOKLAAEAITVORVHGDHLHGLVTRPES  
WLITDFEGEPQPLDERRAPDPLRVAGLRSFEYAAIYPLVDQATDKQLAARREW  
VERNRAAFCDGYAVASGIDPRDSALLLGVELDKAVYETGYETRRHPCWLPILURISA  
RLTAS"

8038..8817

/gene="MT0136"

8038..8817

/gene="MT0136"

/note="identified by Glimmer2; putative"

/codon\_start=1

/transl\_table=11

/product="hypothetical protein"

```
/protein_id="AAK44360.1"
/db_xref="GI:13879618"
/translation="MQREIYDGEARLSWLAAALGILGATFAFTHSAGYFVTFMGNSQ
RAVLGLFGDAMSVTASLLILFFVAGVVIASVCRHFWAAHPHGPTVLTTLSPFAA
GVDIMLGWHEHSMDFPILFVFGVIGALNTSPKDVGEVSVLPYGLYKMGQIE
RHLAGEKYEDWGLGYELLHASFVYLGAAAGGAINSVTGPQNLAVAAVVCAATGITYLH
ADRGVLNKRPPGPKRLFRALRUGELDSGTSTPATNYGSS"
complement(8949..9989)
/gene="MT0137"
/complement(8949..9989)
/gene="MT0137"
/notes="similar to SP:P31953 CB:X57229 PID:48828
PID:1567734 PID:1877254; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="esterase, putative, antigen 85-C"
/protein_id="AAK44361.1"
/db_xref="GI:13879619"
/translation="MGLVAMTFFEOVRRLRSAAATLPRRLAIAAMGAVLVYGLVGT
GGPATAGASRECLPVEYLQVPSASGRDIKVOFGGPHAVYLLDGLRAQDDYNGWD
INTPAEYIQSGLSVIMPVGQSFYTDWYQFOSGONHYTKWETFLREMPAWLQ
ANKGVPTGNAAVLSMVGSGSALILAAVYPOQFPVAAASLGFNLSEGWPTLGLAM
NDSGGVNSAMGSPSSDPKWRNDPMVQIPRLVANNTRVWYCGNGTSPSDLGDDNIPA
KELEGLTLTNTFRTDYAADGRNGVFNPPNGTHSWPINEQLVAMKADIQHVILNG
ATPPAAPAAPAA"
10218..10673
/gene="MT0138"
10218..10673
/gene="MT0138"
/notes="similar to SP:P08634 PID:46222; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="nodulation protein N-related protein"
/protein_id="AAK44362.1"
/db_xref="GI:13879620"
```

```
Query Match
Best Local Similarity 100.0%; Score 396; DB 1; Length 14029;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGGCGGCTCCGATAAATCCAGCTGCTCCAGGTGGCGAGGATTCGCCATTCGGATC 60
Db |||||||
QY 4188 ACGGCGGCTCCGATAAATCCAGCTGCTCCAGGTGGCGAGGATTCGCCATTCGGATC 4247
Db |||||||
QY 61 GGGCAGCGCATGGCGATCGCGGCCAGATCCGATCGGTGGGGGTCAACCCACCGTTTCA 120
Db |||||||
QY 4248 GGGCAGCGCATGGCGATCGCGGCCAGATCCGATCGGTGGGGGTCAACCCACCGTTTCA 4307
Db |||||||
QY 121 ATCGGCGCTACCGCTTCCTCGGCTTGGGTGTCGACAAACGCGGACGCGACGA 180
Db |||||||
QY 4308 ATCGGCGCTACCGCTTCCTCGGCTTGGGTGTCGACAAACGCGGACGCGACGA 4367
Db |||||||
QY 181 GTCCACACGCTGTCGGGAGCGCTCCGGCGGCAAGTCTCGGATCTCCACCGGCGACGTG 240
Db |||||||
QY 4368 GTCCACACGCTGTCGGGAGCGCTCCGGCGGCAAGTCTCGGATCTCCACCGGCGACGTG 4427
Db |||||||
QY 241 ATCACCGCGGTGACGCGCTCCGATCACTCGGCGACCGGATCGGCGACGCGCTTAA 300
Db |||||||
QY 4428 ATCACCGCGGTGACGCGCTCCGATCACTCGGCGACCGGATCGGCGACGCGCTTAA 4487
Db |||||||
QY 301 GGGCATATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTCGGCGGCGACGCGT 360
Db |||||||
QY 4488 GGGCATATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTCGGCGGCGACGCGT 4547
Db |||||||
QY 361 ACAGGGAAGCTACATTTGCCGAGGACCGCCCGGCC 396
Db |||||||
QY 4548 ACAGGGAAGCTACATTTGCCGAGGACCGCCCGGCC 4583
Db |||||||
```

RESULT 6  
BX248334

```
LOCUS
DEFINITION
BX248334 Mycobacterium bovis subsp. bovis AF2122/97 complete genome; segment
1/14.
ACCESSION
BX248334 BX248333
VERSION
BX248334.1 GI:31616762
KEYWORDS
complete genome.
SOURCE
Mycobacterium bovis subsp. bovis AF2122/97
Mycobacterium bovis subsp. bovis AF2122/97
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
1
REFERENCE
Garnier,T., Eiglmeier,K., Camus,J.-C., Medina,N., Mansoor,H.,
Pryor,M., Duthoy,S., Grondin,S., Lacroix,C., Monsemp,C., Simon,S.,
Harris,B., Atkin,R., Doggett,J., Mayes,R., Keating,L.,
Wheeler,P.R., Parkhill,J., Barrell,B.G., Cole,S.T., Gordon,S.V. and
Hewinson,G.
The complete genome sequence of Mycobacterium bovis
Online Publication
PNAS 10.1073/pnas.1130426100 ( Microbiology )
2 (bases 1 to 343050)
REFERENCE
Garnier,T.
Direct Submission
Submitted (24-MAR-2003) Garnier T., Unite de Genetique Moleculaire
Bacterienne Institut Pasteur 28,rue du Dr Roux 75724 PARIS cedex
15, France. e-mail:tgarnier@pasteur.fr Submitted on behalf of the
Mycobacterium bovis sequencing teams, TB Research Group, Veterinary
Laboratories Agency Weybridge, Woodham Lane, New Haw, Addlestone,
Surrey KT15 3NB, UK. Sanger Centre, Wellcome Trust Genome Campus,
Hinxton,Cambridge CB10 1SA, UK. P4 Annotation, Genopole, Institut
Pasteur, 28 Rue du Docteur Roux, 75724 Paris Cedex 15, France.
Unite de Genetique Moleculaire Bacterienne, Institut Pasteur, 28
rue du Docteur Roux, 75724 Paris Cedex 15, France
LOCATION/Qualifiers
1..343050
/organism="Mycobacterium bovis subsp. bovis AF2122/97"
/mol_type="genomic DNA"
/strain="AF2122/97"
/db_xref="taxon:233413"
1..1524
/gene="dnaA"
/locus_tag="Mb0001"
1..1524
/gene="dnaA"
/locus_tag="Mb0001"
/notes="Mb0001, dnaA, len: 507 aa. Equivalent to Rv0001,
len: 507 aa, from Mycobacterium tuberculosis strain H37Rv,
(99.8% identity in 507 aa overlap). dnaA, chromosomal
replication initiator protein (see citations below),
equivalent to other Mycobacterial CHROMOSOMAL REPLICATION
INITIATOR PROTEINS e.g. P46388|DNAA_MYCLE from
Mycobacterium leprae (502 aa); Q9L7L7|DNAA_MYCPA from
Mycobacterium paratuberculosis (509 aa); P49990|DNAA_MYCAV
from Mycobacterium avium (508 aa); P49992|DNAA_MYCSM from
Mycobacterium smegmatis (504 aa); etc. Also highly similar
to others except in N-terminus e.g. Q9ZH75|DNAA_STRCH
CHROMOSOMAL REPLICATION INITIATOR PROTEIN from
Streptomyces chrysomallus (624 aa); Q9ZH76|DNAA_STRRE from
Streptomyces reticuli (643 aa); DNAA_ECOLI|P03004|B3702
chromosomal replication initiator protein from Escherichia
coli strain K12 (467 aa), FASTA scores: opt: 986, E(): 0,
(43.2% identity in 389 aa overlap); etc. Contains P500017
ATP/GTP-binding site motif A (P-loop) and P501008 DnaA
protein signature. BELONGS TO THE DNAA FAMILY. Note that
the first base of this gene has been taken as base 1 of
the Mycobacterium bovis genomic sequence."
/codon_start=1
/evidence=experimental
/transl_table=11
/product="CHROMOSOMAL REPLICATION INITIATOR PROTEIN DNAA"
/protein_id="CAD92863.1"
/db_xref="GI:31616763"
/translation="MTDDPGSGFTTWNNAVVSSELNGDPKVDGPPSSDANLSAPLPQO
```



Db 152188 GTCCAAACGGCTGGTCCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGACGGTG 152247  
QY 241 ATCACCGCGTCCAGCGGCTCCGATCACTACCGCCACCGCGATCGCGACCGCTTAAC 300  
Db 152248 ATCACCGCGTCCAGCGGCTCCGATCACTACCGCCACCGCGATCGCGACCGCTTAAC 152307  
QY 301 GGGCATATCCCGGTGACGTATCTCTCGGTGACCTGGCAACCAAGTCGGCGGCGACGGGT 360  
Db 152308 GGGCATATCCCGGTGACGTATCTCTCGGTGACCTGGCAACCAAGTCGGCGGCGACGGGT 152367  
QY 361 ACAGGGAAGCTGACATTTGGCCGAGGAGACCCCGGCC 396  
Db 152368 ACAGGGAAGCTGACATTTGGCCGAGGAGACCCCGGCC 152403

RESULT 7  
AR169152  
LOCUS AR169152 447 bp DNA linear PAT 17-DEC-2001  
DEFINITION Sequence 4 from patent US 6290969.  
ACCESSION AR169152  
VERSION AR169152.1 GI:17906927  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 447)  
AUTHORS Reed, S.G., Skeiky, Y.A.W., Dillon, D.C., Campos-Neto, A., Houghton, R.,  
Vedvick, T.S. and Twardzik, D.R.  
TITLE Compounds and methods for immunotherapy and diagnosis of  
tuberculosis  
JOURNAL Patent: US 6290969-A 4 18-SEP-2001;  
FEATURES Location/Qualifiers  
source 1. .447  
BASE COUNT 79 a 146 c 149 g 72 t 1 others  
ORIGIN

Query Match 99.6%; Score 394.4; DB 6; Length 447;  
Best Local Similarity 99.7%; Pred. No. 9.2e-57;  
Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACGGCGCGTCCGATCACTCCAGCTGTCACAGGTGGCGAGGATTCGCCATTCGGATC 60  
Db 11 ACGGCGCGTCCGATCACTCCAGCTGTCACAGGTGGCGAGGATTCGCCATTCGGATC 70  
QY 61 GGGCAGGCGATGGCGATCGCGGCCAGATCCGATCGGTTGGGGTTCACCCACCGTTTAT 120  
Db 71 GGGCAGGCGATGGCGATCGCGGCCAGATCCGATCGGTTGGGGTTCACCCACCGTTTAT 130  
QY 121 ATCGGCGCTACCGCTTCTCGGCTTGGGTGTTGTCGACAAACGCAACGGCGCACGA 180  
Db 131 ATCGGCGCTACCGCTTCTCGGCTTGGGTGTTGTCGACAAACGCAACGGCGCACGA 190  
QY 181 GTCCAAACGGTGGTGGGAGCGCTCCGCGGCAAGTCTCGGCATCTCCACCGCGCACGTG 240  
Db 191 GTCCAAACGGTGGTGGGAGCGCTCCGCGGCAAGTCTCGGCATCTCCACCGCGCACGTG 250  
QY 241 ATCACCGCGTCCAGCGGCTCCGATCACTACCGCCACCGCGATCGCGACCGCTTAAC 300  
Db 251 ATCACCGCGTCCAGCGGCTCCGATCACTACCGCCACCGCGATCGCGACCGCTTAAC 310  
QY 301 GGGCATATCCCGGTGACGTATCTCTCGGTGACCTGGCAACCAAGTCGGCGGCGACGGGT 360  
Db 311 GGGCATATCCCGGTGACGTATCTCTCGGTGACCTGGCAACCAAGTCGGCGGCGACGGGT 370  
QY 361 ACAGGGAAGCTGACATTTGGCCGAGGAGACCCCGGCC 396  
Db 371 ACAGGGAAGCTGACATTTGGCCGAGGAGACCCCGGCC 406

RESULT 8  
AR182442  
LOCUS AR182442 447 bp DNA linear PAT 20-APR-2002  
DEFINITION Sequence 4 from patent US 6350456.  
ACCESSION AR182442  
VERSION AR182442.1 GI:20244262  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 447)  
AUTHORS Reed, S.G., Skeiky, Y.A.W. and Dillon, D.C.  
TITLE Compositions and methods for the prevention and treatment of M.  
tuberculosis infection  
JOURNAL Patent: US 6350456-A 4 26-FEB-2002;  
FEATURES Location/Qualifiers  
source 1. .447  
BASE COUNT 79 a 146 c 149 g 72 t 1 others  
ORIGIN

Query Match 99.6%; Score 394.4; DB 6; Length 447;  
Best Local Similarity 99.7%; Pred. No. 9.2e-57;  
Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACGGCGCGTCCGATCACTCCAGCTGTCACAGGTGGCGAGGATTCGCCATTCGGATC 60  
Db 11 ACGGCGCGTCCGATCACTCCAGCTGTCACAGGTGGCGAGGATTCGCCATTCGGATC 70  
QY 61 GGGCAGGCGATGGCGATCGCGGCCAGATCCGATCGGTTGGGGTTCACCCACCGTTTAT 120  
Db 71 GGGCAGGCGATGGCGATCGCGGCCAGATCCGATCGGTTGGGGTTCACCCACCGTTTAT 130  
QY 121 ATCGGCGCTACCGCTTCTCGGCTTGGGTGTTGTCGACAAACGCAACGGCGCACGA 180  
Db 131 ATCGGCGCTACCGCTTCTCGGCTTGGGTGTTGTCGACAAACGCAACGGCGCACGA 190  
QY 181 GTCCAAACGGTGGTGGGAGCGCTCCGCGGCAAGTCTCGGCATCTCCACCGCGCACGTG 240  
Db 191 GTCCAAACGGTGGTGGGAGCGCTCCGCGGCAAGTCTCGGCATCTCCACCGCGCACGTG 250  
QY 241 ATCACCGCGTCCAGCGGCTCCGATCACTACCGCCACCGCGATCGCGACCGCTTAAC 300  
Db 251 ATCACCGCGTCCAGCGGCTCCGATCACTACCGCCACCGCGATCGCGACCGCTTAAC 310  
QY 301 GGGCATATCCCGGTGACGTATCTCTCGGTGACCTGGCAACCAAGTCGGCGGCGACGGGT 360  
Db 311 GGGCATATCCCGGTGACGTATCTCTCGGTGACCTGGCAACCAAGTCGGCGGCGACGGGT 370  
QY 361 ACAGGGAAGCTGACATTTGGCCGAGGAGACCCCGGCC 396  
Db 371 ACAGGGAAGCTGACATTTGGCCGAGGAGACCCCGGCC 406

Best Local Similarity	99.7%	Pred. No. 9.2e-57;								
Matches	395;	Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;	
QY	1	ACGCCGCGTCCGATAACTTTCACAGTCTCCACAGGTGGCAGGATTTCGCATTCGCCATC	60							
Db	11	ACGCCGCGTCCGATAACTTTCACAGTCTCCACAGGTGGCAGGATTTCGCATTCGCCATC	70							
QY	61	GGCAGCGCATCGCATCGCGGCCACAGATCCGATCGGTGGGGGTTCACCCACCGTTTCAT	121							
Db	71	GGCAGCGCATCGCATCGCGGCCACAGATCCGATCGGTGGGGGTTCACCCACCGTTTCAT	130							
QY	121	ATCGGGCTACCGCTTCCTCGGTTCGGTGTGTGTCGACAACAGCGACAGGGCGCACGA	180							
Db	131	ATCGGGCTACCGCTTCCTCGGTTCGGTGTGTGTCGACAACAGCGACAGGGCGCACGA	190							
QY	181	GTCACAGCGTGTGCGGAGCGTCCGGCGCGAAGTCTCGGCATCTCCACCGGGCAGGTG	240							
Db	191	GTCACAGCGTGTGCGGAGCGTCCGGCGCGAAGTCTCGGCATCTCCACCGGGCAGGTG	250							
QY	241	ATCACGCGGTGCACGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGTTTAAAC	300							
Db	251	ATCACGCGGTGCACGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGTTTAAAC	310							
QY	301	GGGCATCATCCCGGTGACGTATCTCGGTGACCTTGGCAAAACCAAGTCGGCGGCACGCGT	360							
Db	311	GGGCATCATCCCGGTGACGTATCTCGGTGACCTTGGCAAAACCAAGTCGGCGGCACGCGT	370							
QY	361	ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC	396							
Db	371	ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC	406							

RESULT 10	AR233097	AR233097	Sequence 4	447 bp	DNA	linear	PAT 20-DEC-2002
LOCUS	AR233097	AR233097	Sequence 4	447 bp	DNA	linear	PAT 20-DEC-2002
DEFINITION	AR233097	AR233097	Sequence 4	447 bp	DNA	linear	PAT 20-DEC-2002
ACCESSION	AR233097	AR233097	Sequence 4	447 bp	DNA	linear	PAT 20-DEC-2002
VERSION	AR233097.1	AR233097.1	Sequence 4	447 bp	DNA	linear	PAT 20-DEC-2002
	GI:27275533	GI:27275533	Sequence 4	447 bp	DNA	linear	PAT 20-DEC-2002

ORGANISM	Unknown.
	Unclassified.
REFERENCE	1 (bases 1 to 447)
AUTHORS	Reed, S.G., Skeiky, Y.A.W., Dillon, D.C., Campos-Neto, A., Houghton, R., Vedvick, T.S., Twardzik, D.R., Lodes, M.J. and Hendrickson, R.C.
TITLE	Compounds and methods for diagnosis of tuberculosis
JOURNAL	Patent: US 6459366-A 4 01-OCT-2002;
FEATURES	Location/Qualifiers
source	1..447

source	79 a	146 c	149 g	72 t	1 others
BASE COUNT					
ORIGIN					
Query Match		99.6%	Score 394.4;	DB 6;	Length 447;
Best Local Similarity		99.7%	Pred. No. 9.2e-57;		
Matches 395;	Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

QY	1	ACGCCCGCTCCGATAACTTCCAGCTGTCCAGGGTGGCAGGGATTCCGCATTTCCGATC	60
Db	11	ACGGCCGCTCCGATAACTTCCAGCTGTCCAGGGTGGCAGGGATTTCGCCATTTCCGATC	70
QY	61	GGCAGCGGATGGCGATCGCGGGGCCAGATCCGATCGGGTGGGGGGTCAACCCACCGTTTCAT	120
Db	71	GGCAGCGGATGGCGATCGCGGGGCCAGATCCGATCGGGTGGGGGGTCAACCCACCGTTTCAT	130
QY	121	ATCGGGGCTACCGCTTCCTCGGCTTGGGTGTTGTCACAAACAGCGCAACGGGGCGACGA	180
Db	131	ATCGGGGCTACCGCTTCCTCGGCTTGGGTGTTGTCACAAACAGCGCAACGGGGCGACGA	190
QY	181	GTCAAACGCTGTGTCGGGAGCGGTTCGGCGGCGAAGTCTCGGCAATCTCCACCGGGCAGGTG	240
Db	191	GTCAAACGCTGTGTCGGGAGCGGTTCGGCGGCGAAGTCTCGGCAATCTCCACCGGGCAGGTG	250

Qy	241	ATCACCGGGTGCAGCGGGTCCGATCAACTCGGCCACCGCGATGCGGGACGCGCTTAAC	300
Db	251	ATCACCGGGTGCAGCGGGTCCGATCAACTCGGCCACCGCGATGCGGGACGCGCTTAAC	310
Qy	301	GGGCATATCCCGGTGACGTATCTCGGTGACCTTGGCAAAACCAAGTCGGGGCGGCACGGT	360
Db	311	GGGCATATCCCGGTGACGTATCTCGGTGAACTTGGCAAAACCAAGTCGGGGCGGCAGGT	370
Qy	361	ACAGGAAACGTGACATTGGCGAGGGACCCCGGCC	396
Db	371	ACAGGAAACGTGACATTGGCGAGGGACCCCGGCC	406

RESULT 11	AX429596	AX429596	AX429596	Sequence 4 from Patent EP1203817.	447 bp	DNA	linear	PAT 21-JUN-2002
LOCUS	AX429596	AX429596	AX429596	Sequence 4 from Patent EP1203817.	447 bp	DNA	linear	PAT 21-JUN-2002
DEFINITION	AX429596	AX429596	AX429596	Sequence 4 from Patent EP1203817.	447 bp	DNA	linear	PAT 21-JUN-2002
ACCESSION	AX429596	AX429596	AX429596	Sequence 4 from Patent EP1203817.	447 bp	DNA	linear	PAT 21-JUN-2002
VERSION	AX429596.1	AX429596.1	AX429596.1	Sequence 4 from Patent EP1203817.	447 bp	DNA	linear	PAT 21-JUN-2002
KEYWORDS	AX429596.1	AX429596.1	AX429596.1	Sequence 4 from Patent EP1203817.	447 bp	DNA	linear	PAT 21-JUN-2002
SOURCE	AX429596.1	AX429596.1	AX429596.1	Sequence 4 from Patent EP1203817.	447 bp	DNA	linear	PAT 21-JUN-2002
ORGANISM	AX429596.1	AX429596.1	AX429596.1	Sequence 4 from Patent EP1203817.	447 bp	DNA	linear	PAT 21-JUN-2002

REFERENCE	1
AUTHORS	Reed, S.G., Skeiky, Y.A., Dillon, D.C., Campos-Neto, A., Houghton, R.L., Vedvick, T.S. and Twardzik, D.R.
TITLE	Compounds and methods for immunotherapy and diagnosis of tuberculosis
JOURNAL	Patent: EP 1203817-A 4 08-MAY-2002; CORIXA CORPORATION (US)
FEATURES	Location/Qualifiers
SOURCE	1. .447

BASE COUNT	79 a	146 c	149 g	72 t	1 others
ORIGIN					
Query Match		99.6%	Score 394.4;	DB 6;	Length 447;
Best Local Similarity		99.7%;	Pred. No. 9.2e-57;		
Matches 395;		Conservative	0;	Mismatches 1;	Indels 0;
Gaps 0;					

Qy	1	ACGGCCGGTCCGATAACTTCCAGGTGTCCAGGGTGGCAGGATTCGCCATTCGGATC	60
Db	11	ACGGCCGGTCCGATAACTTCCAGGTGTCCAGGGTGGCAGGATTCGCCATTCGGATC	70
Qy	61	GGGCAGCGATGGGATCGCGGGCCAGATCCGATCGGTGGGTGCACCCACACGTTTCAT	120
Db	71	GGGCAGCGATGGGATCGCGGGCCAGATCCGATCGGTGGGTGCACCCACGCTTCAT	130
Qy	121	ATCGGGCTACCGCTTCTCGGCTTGGGTCTTGTCCACAACACGGCAACGGGCGACGA	180
Db	131	ATCGGGCTACCGCTTCTCGGCTTGGGTCTTGTCCACAACACGGCAACGGGCGACGA	190
Qy	181	GTCACACGGTGTGGGAGCGCTTCGGGGCAAGTCTCGGCATCTCACCGGGGACGTG	240
Db	191	GTCACACGGTGTGGGAGCGCTTCGGGGCAAGTCTCGGCATCTCACCGGGGACGTG	250
Qy	241	ATCACCGGTGCAGCGGCTCCGATCAACTCGGCCACCGGATGGGGAGCGGTTAAAC	300
Db	251	ATCACCGGTGCAGCGGCTCCGATCAACTCGGCCACCGGATGGGGAGCGGTTAAAC	310
Qy	301	GGGCATATCCCGTGAAGTCTCTCGGTGACCTTGGCAACCAAGTCGGGGGACGGGT	360
Db	311	GGGCATATCCCGTGAAGTCTCTCGGTGACCTTGGCAACCAAGTCGGGGGACGGGT	370
Qy	361	ACAGGGAACGTGACATTGGCCGAGGGACCCCGGCC	396
Db	371	ACAGGGAACGTGACATTGGCCGAGGGACCCCGGCC	406

RESULT 12  
BD006325

LOCUS BD006325 447 bp DNA linear PAT 31-JAN-2002  
 DEFINITION Compounds and methods for diagnosis of Tuberculosis.  
 ACCESSION BD006325  
 VERSION BD006325.1 GI:18634696  
 KEYWORDS JP 2001500383-A/4.  
 SOURCE unclassified  
 ORGANISM unclassified  
 REFERENCE 1 (bases 1 to 447)  
 AUTHORS Reed, S.G., Skeiky, Y.A.W., Dillon, D.C., Neto, A.C., Houghton, R.,  
 Vedvick, T.S., Twardzik, D.R., and Lodes, M.J.  
 TITLE Compounds and methods for diagnosis of Tuberculosis  
 JOURNAL Patent: JP 2001500383-A 4 16-JAN-2001;  
 CORIXA CORP  
 COMMENT OS Unidentified  
 PN JP 2001500383-A/4  
 PD 16-JAN-2001  
 PF 07-OCT-1997 JP 1998518432  
 PR 11-OCT-1996 US 08/729622,13-MAR-1997 US 08/818111 PI  
 STEVEN G REED, YASIR A W SKEIKY, DAVIN C DILLON, PI ANTONIO CAMPOS  
 NETO,  
 PI RAYMOND HOUGHTON, THOMAS S VEDVICK, DANIEL R TWARDZIK, PI  
 MICHAEL J LODES  
 PC C12N15/31, C07K14/35, C07K16/12, C12Q1/68, C12N15/62, C01N33/53 CC  
 Strandedness: Single;  
 CC Topology: Linear;  
 FH Key Location/Qualifiers  
 FT source 1..447 /organism='Unidentified'.  
 FEATURES  
 source  
 1..447 Location/Qualifiers  
 /organism='unidentified'  
 /mol\_type='genomic DNA'  
 /db\_xref='taxon:32644'  
 BASE COUNT 79 a 146 c 149 g 72 t 1 others  
 ORIGIN  
 Query Match 99.6%; Score 394.4; DB 6; Length 447;  
 Best Local Similarity 99.7%; Pred. No. 9.2e-57;  
 Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 ACGGCGGCTCCGATAACTTCCAGCTGTCCAGGCTGGGAGGATTCGCCATTCCGATC 60  
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 11 ACGGCGGCTCCGATAACTTCCAGCTGTCCAGGCTGGGAGGATTCGCCATTCCGATC 70  
 QY 61 GGGCAGGCGATGGCGATCGCGGCCAGATCCGATCGGTGGGGTTCACCCACCGTTTAT 120  
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 71 GGGCAGGCGATGGCGATCGCGGCCAGATCCGATCGGTGGGGTTCACCCACCGTTTAT 130  
 QY 121 ATCGGCGCTACCGCTTCTCGGCTTGGGTGTTGTCGACAAACGCGGACGCGT 180  
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 131 ATCGGCGCTACCGCTTCTCGGCTTGGGTGTTGTCGACAAACGCGGACGCGT 190  
 QY 181 GTCCAAACGCGTGGTGGGAGCGCTCCGCGGCAAGTCTCGGATCTCCACCGGCGACGTG 240  
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 191 GTCCAAACGCGTGGTGGGAGCGCTCCGCGGCAAGTCTCGGATCTCCACCGGCGACGTG 250  
 QY 241 ATCACCAGGCGTGGTGGGAGCGCTCCGCGGCAAGTCTCGGATCTCCACCGGCGACGTG 300  
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 251 ATCACCAGGCGTGGTGGGAGCGCTCCGCGGCAAGTCTCGGATCTCCACCGGCGACGTG 310  
 QY 301 GGGCATATCCGGGTGAGCTCATCTCGGTGAGCTGGCAACCAAGTTCGGGCGGCGT 360  
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 311 GGGCATATCCGGGTGAGCTCATCTCGGTGAGCTGGCAACCAAGTTCGGGCGGCGT 370  
 QY 361 ACAGGAACTGACATTTGGCGGAGGACCCCGGCC 396  
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 371 ACAGGAACTGACATTTGGCGGAGGACCCCGGCC 406

RESULT 13  
 BD006445  
 LOCUS

DEFINITION Compounds and methods for immunotherapy and diagnosis of  
 Tuberculosis.  
 ACCESSION BD006445  
 VERSION BD006445.1 GI:18634816  
 KEYWORDS JP 2001501832-A/4.  
 SOURCE unclassified  
 ORGANISM unclassified  
 REFERENCE 1 (bases 1 to 447)  
 AUTHORS Reed, S.G., Skeiky, Y.A.W., Dillon, D.C., Neto, A.C., Houghton, R.,  
 Vedvick, T.S., Twardzik, D.R., and Lodes, M.J.  
 TITLE Compounds and methods for immunotherapy and diagnosis of  
 JOURNAL Patent: JP 2001501832-A 4 13-FEB-2001;  
 CORIXA CORP  
 COMMENT OS Unidentified  
 PN JP 2001501832-A/4  
 PD 13-FEB-2001  
 PF 07-OCT-1997 JP 1998518456  
 PR 11-OCT-1996 US 08/730510,13-MAR-1997 US 08/818112 PI  
 STEVEN G REED, YASIR A W SKEIKY, DAVIN C DILLON, PI ANTONIO CAMPOS  
 NETO,  
 PI RAYMOND HOUGHTON, THOMAS S VEDVICK, DANIEL R TWARDZIK, PI  
 MICHAEL J LODES  
 PC C12N15/31, C07K14/35, A61K39/04, A61K48/00, A61K49/00, C12N15/62,  
 C07K19/00,  
 PC G01N33/50, G01N33/60, G01N33/569, C12N1/19, C12N1/20, C12N1/21, PC  
 C12N5/10//  
 PC (C12N1/21, C12N1/19)  
 CC Strandedness: Single;  
 CC Topology: Linear;  
 FH Key Location/Qualifiers  
 FT source 1..447 /organism='Unidentified'.  
 FEATURES  
 source  
 1..447 Location/Qualifiers  
 /organism='unidentified'  
 /mol\_type='genomic DNA'  
 /db\_xref='taxon:32644'  
 BASE COUNT 79 a 146 c 149 g 72 t 1 others  
 ORIGIN  
 Query Match 99.6%; Score 394.4; DB 6; Length 447;  
 Best Local Similarity 99.7%; Pred. No. 9.2e-57;  
 Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 ACGGCGGCTCCGATAACTTCCAGCTGTCCAGGCTGGGAGGATTCGCCATTCCGATC 60  
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 11 ACGGCGGCTCCGATAACTTCCAGCTGTCCAGGCTGGGAGGATTCGCCATTCCGATC 70  
 QY 61 GGGCAGGCGATGGCGATCGCGGCCAGATCCGATCGGTGGGGTTCACCCACCGTTTAT 120  
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 71 GGGCAGGCGATGGCGATCGCGGCCAGATCCGATCGGTGGGGTTCACCCACCGTTTAT 130  
 QY 121 ATCGGCGCTACCGCTTCTCGGCTTGGGTGTTGTCGACAAACGCGGACGCGT 180  
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 131 ATCGGCGCTACCGCTTCTCGGCTTGGGTGTTGTCGACAAACGCGGACGCGT 190  
 QY 181 GTCCAAACGCGTGGTGGGAGCGCTCCGCGGCAAGTCTCGGATCTCCACCGGCGACGTG 240  
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 191 GTCCAAACGCGTGGTGGGAGCGCTCCGCGGCAAGTCTCGGATCTCCACCGGCGACGTG 250  
 QY 241 ATCACCAGGCGTGGTGGGAGCGCTCCGCGGCAAGTCTCGGATCTCCACCGGCGACGTG 300  
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 251 ATCACCAGGCGTGGTGGGAGCGCTCCGCGGCAAGTCTCGGATCTCCACCGGCGACGTG 310  
 QY 301 GGGCATATCCGGGTGAGCTCATCTCGGTGAGCTGGCAACCAAGTTCGGGCGGCGT 360  
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 311 GGGCATATCCGGGTGAGCTCATCTCGGTGAGCTGGCAACCAAGTTCGGGCGGCGT 370  
 QY 361 ACAGGAACTGACATTTGGCGGAGGACCCCGGCC 396  
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 371 ACAGGAACTGACATTTGGCGGAGGACCCCGGCC 406

```
RESULT 14
BD069285
LOCUS
DEFINITION
  BD069285
  Compounds and methods for immunotherapy and diagnosis of
  tuberculosis.
ACCESSION
  BD069285
VERSION
  JP 2001517069-A/4.
KEYWORDS
  unidentified
SOURCE
  unidentified
ORGANISM
  unclassified.
REFERENCE
  1 (bases 1 to 447)
  Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Neto,A.C., Houghton,R.,
  Vedvick,T.H. and Twardzik,D.R.
  Compounds and methods for immunotherapy and diagnosis of
  tuberculosis.
  Patent: JP 2001517069-A 4 02-OCT-2001;
  CORIXA CORP
  OS Unidentified
  PN JP 2001517069-A/4
  PD 02-OCT-2001
  PR 30-AUG-1996 JP 1997511464
  PR 01-SEP-1995 US 08/523436,22-SEP-1995 US 08/533634 PR
  22-MAR-1996 US 08/620874,05-JUN-1996 US 08/659683 PR
  12-JUL-1996 US 08/680574
  PI STEVEN G REED,YASIR A W SKEIKY,DAVIN C DILLON,ANTONIO CAMPOS
  PI NETO,
  PI RAYMOND HOUGHTON,THOMAS H VEDVICK,DANIEL R TWARDZIK PC
  C12N15/31,C07K14/35,A61K38/16,C12N15/62,G01N33/569,C12Q1/68, PC
  C12N5/10,
  PC C12N1/21/A61K39/04,(C12N1/21,C12R1:19)
  CC Strandedness: Single;
  CC Topology: Linear;
  CC Compounds and methods for immunotherapy and diagnosis of CC
  tuberculosis
  FH Key Location/Qualifiers
  FT source 1..447
  FT /organism='Unidentified'.
  FEATURES
    source
      Location/Qualifiers
        1..447
          /organism='unidentified'
          /mol_type='genomic DNA'
          /db_xref='taxon:32644'
  BASE COUNT
    79 a 146 c 149 g 72 t 1 others
  ORIGIN
    Query Match 99.68; Score 394.4; DB 6; Length 447;
    Best Local Similarity 99.7%; Pred. No. 9.2e-57;
    Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
  Qy 1 ACGGCCGCTCCGATTAACCTCCAGCTGTCCAGGGTGGGAGGATTCGGCATTCGGATC 60
  Db 11 ACGGCCGCTCCGATTAACCTCCAGCTGTCCAGGGTGGGAGGATTCGGCATTCGGATC 70
  Qy 61 GGGCAGCGATGGCGATCGGGGCGAGATCCGATCGGATGGGGGTGACCCACCGTTTCAT 120
  Db 71 GGGCAGCGATGGCGATCGGGGCGAGATCCGATCGGATGGGGGTGACCCACCGTTTCAT 130
  Qy 121 ATCGGGCTACCGCTTCCCTCGGCTTGGGTGTTGTCGACACACGCGGACGGGACGGA 180
  Db 131 ATCGGGCTACCGCTTCCCTCGGCTTGGGTGTTGTCGACACACGCGGACGGGACGGA 190
  Qy 181 GTCAACCGCTGGTTCGGGAGCGTCCGGCGGCAAGTCTCGGCATCTCCACCGGACGTG 240
  Db 191 GTCAACCGCTGGTTCGGGAGCGTCCGGCGGCAAGTCTCGGCATCTCCACCGGACGTG 250
  Qy 241 ATCAACCGCTGGTTCGGGAGCGTTCGGATCAACTCGGCCACCGGATGGCGGACGGCTTAA 300
  Db 251 ATCAACCGCTGGTTCGGGAGCGTTCGGATCAACTCGGCCACCGGATGGCGGACGGCTTAA 310
  Qy 301 GGGCATCATCCCGTACGCTCATCTCGGTACCTGGGCAACCAAGTCCGGGCGGACGCGT 360
  Db 311 GGGCATCATCCCGTACGCTCATCTCGGTACCTGGGCAACCAAGTCCGGGCGGACGCGT 370
```

```
Qy 361 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 396
Db 371 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 406
RESULT 15
ARI169165
LOCUS
DEFINITION
  ARI169165
  Sequence 17 from patent US 6290969.
ACCESSION
  ARI169165
VERSION
  ARI169165.1 GI:17906947
KEYWORDS
  SOURCE
  ORGANISM
  Unknown.
  Unclassified.
  1 (bases 1 to 1872)
  Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Campos-Neto,A., Houghton,R.,
  Vedvick,T.S. and Twardzik,D.R.
  Compounds and methods for immunotherapy and diagnosis of
  tuberculosis
  Patent: US 6290969-A 17 18-SEP-2001;
  JOURNAL
  Location/Qualifiers
  FEATURES
    source
      1..1872
        /organism='unknown'
  BASE COUNT
    318 a 616 c 604 g 332 t 2 others
  ORIGIN
    Query Match 98.8%; Score 391.2; DB 6; Length 1872;
    Best Local Similarity 99.2%; Pred. No. 2.4e-56;
    Matches 393; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
  Qy 1 ACGGCCGCTCCGATTAACCTCCAGCTGTCCAGGGTGGGAGGATTCGGCATTCGGATC 60
  Db 758 ACGGCCGCTCCGATTAACCTCCAGCTGTCCAGGGTGGGAGGATTCGGCATTCGGATC 817
  Qy 61 GGGCAGCGATGGCGATCGGGGCGAGATCCGATCGGATGGGGGTGACCCACCGTTTCAT 120
  Db 818 GGGCAGCGATGGCGATCGGGGCGAGATCCGATCGGATGGGGGTGACCCACCGTTTCAT 877
  Qy 121 ATCGGGCTACCGCTTCCCTCGGCTTGGGTGTTGTCGACACACGCGGACGGGACGGA 180
  Db 878 ATCGGGCTACCGCTTCCCTCGGCTTGGGTGTTGTCGACACACGCGGACGGGACGGA 937
  Qy 181 GTCAACCGCTGGTTCGGGAGCGTCCGGCGGCAAGTCTCGGCATCTCCACCGGACGTG 240
  Db 938 GTCAACCGCTGGTTCGGGAGCGTCCGGCGGCAAGTCTCGGCATCTCCACCGGACGTG 997
  Qy 241 ATCAACCGCTGGTTCGGGAGCGTTCGGATCAACTCGGCCACCGGATGGCGGACGGCTTAA 300
  Db 998 ATCAACCGCTGGTTCGGGAGCGTTCGGATCAACTCGGCCACCGGATGGCGGACGGCTTAA 1057
  Qy 301 GGGCATCATCCCGTACGCTCATCTCGGTACCTGGGCAACCAAGTCCGGGCGGACGCGT 360
  Db 1058 GGGCATCATCCCGTACGCTCATCTCGGTACCTGGGCAACCAAGTCCGGGCGGACGCGT 1117
  Qy 361 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 396
  Db 1118 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 1153
  Search completed: September 5, 2003, 07:52:35
  Job time : 1680 secs
```

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2003, 07:20:00 : Search time 246 Seconds  
(without alignments)  
4345.437 Million cell updates/sec

Title: US-09-684-215A-3

Perfect score: 396

Sequence: 1 acgcccgcgtccgataactt.....tgccgagggaccccgcc 396

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_19Jun03.\*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*
- 25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	396	100.0	396	23	NA1980
2	396	100.0	672	23	NA1981
3	396	100.0	702	23	NA20206
4	396	100.0	702	23	NA140770
5	396	100.0	702	24	ABK14140
6	396	100.0	1002	24	AA47077
7	396	100.0	1002	24	AA47078
8	396	100.0	1002	24	AA28336

9	396	100.0	1002	24	AA28337	Mycobacterium spec
10	396	100.0	1068	20	AA34251	Mycobacterium spec
11	396	100.0	1143	20	AA34252	Mycobacterium spec
12	396	100.0	1742	23	AA40771	Nucleotide sequenc
13	396	100.0	2190	24	AA47084	Mycobacterium sp.
14	396	100.0	2190	24	AA28343	Mycobacterium spec
15	396	100.0	2191	23	AA40773	Nucleotide sequenc
16	396	100.0	2286	24	ABK14128	DNA encoding antig
17	396	100.0	2287	20	AA220194	Mycobacterium tube
18	396	100.0	2287	20	AA47083	Mycobacterium spec
19	396	100.0	2287	24	AA28342	Mycobacterium spec
20	396	100.0	2808	24	AA47110	Mycobacterium sp.
21	396	100.0	4403765	22	AA199683	Mycobacterium tube
22	396	100.0	4411529	22	AA199682	Mycobacterium tube
23	394.4	99.6	447	18	AA791466	Mycobacterium tube
24	394.4	99.6	447	18	AA791403	Mycobacterium tube
25	394.4	99.6	447	19	AAV64450	M. tuberculosis im
26	394.4	99.6	447	19	AAV64450	M. tuberculosis im
27	394.4	99.6	447	20	AA219252	M. tuberculosis an
28	394.4	99.6	447	20	AA219040	M. tuberculosis re
29	394.4	99.6	447	22	AA503780	M. tuberculosis DN
30	394.4	99.6	447	24	AA47080	Mycobacterium tube
31	394.4	99.6	447	24	AA28339	Mycobacterium spec
32	394.4	99.6	1629	22	AA503793	M. tuberculosis DN
33	392.8	99.2	1871	23	AA40768	Nucleotide sequenc
34	391.2	98.8	1872	18	AA791477	Mycobacterium tube
35	391.2	98.8	1872	18	AA791414	Mycobacterium tube
36	391.2	98.8	1872	19	AAV64463	M. tuberculosis im
37	391.2	98.8	1872	19	AAV44355	Mycobacterium tube
38	391.2	98.8	1872	20	AA219265	M. tuberculosis an
39	391.2	98.8	1872	20	AA219053	M. tuberculosis re
40	391.2	98.8	1872	22	AA503781	M. tuberculosis DN
41	391.2	98.8	1872	24	AA47076	Mycobacterium tube
42	391.2	98.8	1872	24	AA28335	Mycobacterium spec
43	384	97.0	1067	22	AA51961	Mycobacterium tube
44	354	89.4	675	22	AA564132	Human /M. tubercul
45	354	89.4	675	22	AAH93896	Ral2-P510S-C const

## ALIGNMENTS

### RESULT 1

AA40769  
ID AA40769 standard; DNA; 396 BP.

XX  
AC AA40769;

XX  
DT 03-OCT-2002 (first entry)

XX  
DE Nucleotide sequence encoding Ral2 protein.

XX  
KW Ral2; serine protease antigen; MTB32A; Mycobacterium tuberculosis;  
vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;  
immunogen; cytokine; gene; ds.

XX  
OS Unidentified.

XX  
FH Key Location/Qualifiers

FT CDS 1..396

FT /\*tag= a

FT /product= "Ral2 protein"

FT /note= "No start or stop codon"

XX  
PN W0200125401-A2.

XX  
PD 12-APR-2001.

XX  
PF 06-OCT-2000; 2000WO-US27652.

XX  
PR 07-OCT-1999; 99US-0158585.

XX  
PA (CORI-) CORIXA CORP.

```
XX PI Skeiky Y, Guderian J;
XX
XX WPI; 2001-266299/27.
DR P-PSDB; AAO22138.
XX
XX Recombinant nucleic acid molecule for producing high yield expression
PT of desired fusion polypeptides, encodes fusion polypeptide comprising
PT Mycobacterium tuberculosis coding sequence and heterologous polypeptide
PT
XX
XX Claim 1; Fig 2; 39pp; English.
XX
XX The invention relates to a recombinant nucleic acid molecule encoding a
CC fusion polypeptide, comprising a polynucleotide sequence of Ral2, a 14
CC kDa C-terminal fragment of serine protease antigen MTB32A of
CC Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.
CC The recombinant fusion nucleic acids and polypeptides are useful for
CC providing stable and high yield expression of fusion polypeptides of both
CC eukaryotic and prokaryotic origin and to encode a protein product for use
CC as an antigen for detecting serum antibodies. The presence of serum
CC antibodies to M. tuberculosis antigens in an individual indicates that
CC the individual is infected with it. The fusion polypeptides are useful as
CC sources of proteins for monitoring binding of serum antibodies to fusion
CC proteins and as an immunogen to induce and/or enhance immune responses.
CC The coding sequences can be ligated with a coding sequence of another
CC molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and
CC can be used in vivo as a DNA vaccine. This polynucleotide sequence
CC represents the DNA encoding the Ral2 protein.
XX
XX Sequence 396 BP; 58 A; 132 C; 134 G; 52 T; 0 other;
XX
XX Query Match 100.0%; Score 396; DB 23; Length 396;
XX Best Local Similarity 100.0%; Pred. No. 1.1e-84;
XX Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 ACGGCGGCTCCGATAACTTCCAGCTGTCCAGGTGGCAGGATTCGCCATTCGATC 60
Db 1 ACGGCGGCTCCGATAACTTCCAGCTGTCCAGGTGGCAGGATTCGCCATTCGATC 60
QY 61 GGGCAGGCGATGGCGATCGCGGCGCAGATCCGATCGGTGGGGTTCACCCACCGTTTAT 120
Db 61 GGGCAGGCGATGGCGATCGCGGCGCAGATCCGATCGGTGGGGTTCACCCACCGTTTAT 120
QY 121 ATCGGGCTTACCGCTTCTCGGCTTGGGTGTTGTCGACAAACGGAACGGCGCACGA 180
Db 121 ATCGGGCTTACCGCTTCTCGGCTTGGGTGTTGTCGACAAACGGAACGGCGCACGA 180
QY 181 GTCCAAACGGTGGTGGGAGCGCTCGGGGCAAGTCTCGGCATCTCCACGGCGACGTG 240
Db 181 GTCCAAACGGTGGTGGGAGCGCTCGGGGCAAGTCTCGGCATCTCCACGGCGACGTG 240
QY 241 ATACCGGGGTGACGGGCTCCGATCACTCGGCACCGCGATGGCGGACGCGCTTAAC 300
Db 241 ATACCGGGGTGACGGGCTCCGATCACTCGGCACCGCGATGGCGGACGCGCTTAAC 300
QY 301 GGGCATATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTCGGGGGCGACGGT 360
Db 301 GGGCATATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTCGGGGGCGACGGT 360
QY 361 ACAGGGAAGGTGACATTTGGCCGAGGACCCCGGCC 396
Db 361 ACAGGGAAGGTGACATTTGGCCGAGGACCCCGGCC 396

RESULT 2
ID AAL40772
XX AAL40772 standard; DNA; 672 BP.
XX
XX AAL40772;
XX
XX 03-OCT-2002 (first entry)
XX
```

```
DE XX Nucleotide sequence encoding Ral2-mammaglobin fusion protein.
KW Ral2; serine protease antigen; MTB32A; Mycobacterium tuberculosis;
KW vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;
KW immunogen; cytokine; gene; ds.
XX
XX Chimeric - Mammalian.
OS Chimeric - Unidentified.
XX
XX Key Location/Qualifiers
FT 4..566
FT /*tag= a
FT /product= "Ral2-mammaglobin fusion protein"
XX
XX WO200125401-A2.
XX
XX 12-APR-2001.
XX
XX 06-OCT-2000; 200WO-US27652.
XX
XX 07-OCT-1999; 99US-0158585.
XX
XX (CORI-) CORIXA CORP.
XX
XX Skeiky Y, Guderian J;
XX
XX WPI; 2001-266299/27.
XX P-PSDB; AAO22141.
XX
XX Recombinant nucleic acid molecule for producing high yield expression
XX of desired fusion polypeptides, encodes fusion polypeptide comprising
XX Mycobacterium tuberculosis coding sequence and heterologous polypeptide
XX
XX Disclosure; Fig 5; 39pp; English.
XX
XX The invention relates to a recombinant nucleic acid molecule encoding a
XX fusion polypeptide, comprising a polynucleotide sequence of Ral2, a 14
XX kDa C-terminal fragment of serine protease antigen MTB32A of
XX Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.
XX The recombinant fusion nucleic acids and polypeptides are useful for
XX providing stable and high yield expression of fusion polypeptides of both
XX eukaryotic and prokaryotic origin and to encode a protein product for use
XX as an antigen for detecting serum antibodies. The presence of serum
XX antibodies to M. tuberculosis antigens in an individual indicates that
XX the individual is infected with it. The fusion polypeptides are useful as
XX sources of proteins for monitoring binding of serum antibodies to fusion
XX proteins and as an immunogen to induce and/or enhance immune responses.
XX The coding sequences can be ligated with a coding sequence of another
XX molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and
XX can be used in vivo as a DNA vaccine. This polynucleotide sequence
XX represents the DNA encoding the Ral2-mammaglobin fusion protein.
XX
XX Sequence 672 BP; 160 A; 185 C; 187 G; 140 T; 0 other;
XX
XX Query Match 100.0%; Score 396; DB 23; Length 672;
XX Best Local Similarity 100.0%; Pred. No. 1.2e-84;
XX Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 ACGGCGGCTCCGATAACTTCCAGCTGTCCAGGTGGCAGGATTCGCCATTCGATC 60
Db 25 ACGGCGGCTCCGATAACTTCCAGCTGTCCAGGTGGCAGGATTCGCCATTCGATC 84
QY 61 GGGCAGGCGATGGCGATCGCGGCGCAGATCCGATCGGTGGGGTTCACCCACCGTTTAT 120
Db 85 GGGCAGGCGATGGCGATCGCGGCGCAGATCCGATCGGTGGGGTTCACCCACCGTTTAT 144
QY 121 ATCGGGCTTACCGCTTCTCGGCTTGGGTGTTGTCGACAAACGGAACGGCGCACGA 180
Db 145 ATCGGGCTTACCGCTTCTCGGCTTGGGTGTTGTCGACAAACGGAACGGCGCACGA 204
QY 181 GTCCAAACGGTGGTGGGAGCGCTCGGGGCAAGTCTCGGCATCTCCACGGCGACGTG 240
Db 181 GTCCAAACGGTGGTGGGAGCGCTCGGGGCAAGTCTCGGCATCTCCACGGCGACGTG 240
```

Db 205 GTCCAAACCGGTGGTGGGAGCGCTCCGGCGGCAAGTCTTCGGCATCTCCACCGGCGACGTG 264  
QY 241 ATCACCGCGGTGCGACGGCGCTCCGATCAACTCGGCCACCGGATGGCGGACGCGCTTAAC 300  
Db 265 ATCACCGGGTTCGACGGGCTCCGATCAACTCGGCCACCGGATGGCGGACGCGCTTAAC 324  
QY 301 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGGCAACAAAGTCGGGCGGACGCGT 360  
Db 325 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGGCAACAAAGTCGGGCGGACGCGT 384  
QY 361 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 396  
Db 385 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 420

## RESULT 3

AAZ20206  
ID AAZ20206 standard; DNA; 702 BP.

XX AC AAZ20206;

XX DT 17-JAN-2000 (first entry)

XX DE Mycobacterium tuberculosis antigen fusion protein Mtb24 DNA.  
XX KW Tuberculosis; antigen; fusion protein; Mtb24; Ra12; DPPD;

XX KW diagnosis; therapy; vaccine; immunogen; ss.  
XX OS Mycobacterium tuberculosis.

PN WO9951748-A2.  
XX PD 14-OCT-1999.

XX PF 07-APR-1999; 99WO-US07717.  
XX PR 07-APR-1998; 98US-0056556.  
XX PR 30-DEC-1998; 98US-0223040.

XX PA (CORI-) CORIXA CORP.  
XX PI Skeiky YAW, Alderson M, Campos-Neto A;

XX DR WPI; 1999-601610/51.  
XX DR P-PSDB; AAY32071.

XX PT New fusion proteins useful for diagnosis, prevention and treatment of tuberculosis -

XX PS Example; Fig 13A-B; 83pp; English.

XX CC This DNA sequence includes a coding region for a recombinant Mycobacterium tuberculosis bi-antigen fusion protein (see AAY32071), termed Mtb24, composed of the antigens Ra12 and DPPD. The DNA is useful for the recombinant production of the fusion protein. Coding sequences for the antigens were modified by PCR in order to facilitate their fusion and subsequent expression of the fusion protein, and then ligated. The invention provides fusion proteins (see AAY32059-71) containing at least 2 M. tuberculosis antigens. The new fusion proteins and polynucleotides encoding them are useful as vaccines for preventing tuberculosis (claimed), for diagnosis (via in vitro assays or intradermal skin tests for detection of anti-M. tuberculosis antibodies), monitoring of disease progression, and treatment of tuberculosis. They are more effective immunogens than mixtures of the individual protein components.

XX SQ Sequence 702 BP; 127 A; 229 C; 225 G; 121 T; 0 other;

Query Match 100.0%; Score 396; DB 20; Length 702;  
Best Local Similarity 100.0%; Pred. No. 1.2e-84;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY i ACGGCGCGTCCGATACTTCACAGTGTCCACGGGTGGGAGGATTGCGCCATTCCGATC 60

Db 25 ACGGCGCGTCCGATACTTCACAGTGTCCACGGGTGGGAGGATTGCGCCATTCCGATC 84  
QY 61 GGGCAGCGGATGGCGATCGCGGCGCAGATCCGATCGGTGGGGTCAACCGGTTTCAT 120  
Db 85 GGGCAGCGGATGGCGATCGCGGCGCAGATCCGATCGGTGGGGTCAACCGGTTTCAT 144  
QY 121 ATCGGGCTACCGGCTTCTCGGCTTGGGTGTTGTCGACAAACAAACGGCAACGGCGCACGA 180  
Db 145 ATCGGGCTACCGGCTTCTCGGCTTGGGTGTTGTCGACAAACAAACGGCAACGGCGCACGA 204  
QY 181 GTCCAAACGCTGTCGGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACGGGCGACGTG 240  
Db 205 GTCCAAACGCTGTCGGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACGGGCGACGTG 264  
QY 241 ATCACCGCGGTGCGACGGGCTCCGATCAACTCGGCCACCGGATGGCGGACGCGTTAAC 300  
Db 265 ATCACCGCGGTGCGACGGGCTCCGATCAACTCGGCCACCGGATGGCGGACGCGTTAAC 324  
QY 301 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGGCAACAAAGTCGGGCGGACGCGT 360  
Db 325 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGGCAACAAAGTCGGGCGGACGCGT 384  
QY 361 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 396  
Db 385 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 420

## RESULT 4

AAAL0770  
ID AAL40770 standard; DNA; 702 BP.

XX AC AAL40770;

XX DT 03-OCT-2002 (first entry)

XX DE Nucleotide sequence encoding Ra12-DPPD fusion protein.

XX KW Ra12; serine protease antigen; Mtb32A; Mycobacterium tuberculosis;  
XX KW vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;  
XX KW immunogen; cytokine; gene; ds.

XX OS Chimeric - Unidentified.

XX FH Location/Qualifiers  
XX CDS 4..696  
XX FT /\*tag= a  
XX FT /product= "Ra12-DPPD fusion protein"

XX PN WO200125401-A2.

XX PD 12-APR-2001.

XX PF 06-OCT-2000; 2000WO-US27652.

XX PR 07-OCT-1999; 99US-0158585.

XX PA (CORI-) CORIXA CORP.

XX PI Skeiky Y, Guderian J;

XX DR WPI; 2001-266299/27.

XX DR P-PSDB; AAO22139.

XX PT Recombinant nucleic acid molecule for producing high yield expression of desired fusion polypeptides, encodes fusion polypeptide comprising Mycobacterium tuberculosis coding sequence and heterologous polypeptide

XX PS Example 1; Fig 3; 39pp; English.

XX CC The invention relates to a recombinant nucleic acid molecule encoding a fusion polypeptide, comprising a polynucleotide sequence of Ra12, a 14

CC kDa C-terminal fragment of serine protease antigen MTB32A of  
 CC Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.  
 CC The recombinant fusion nucleic acids and polypeptides are useful for  
 CC providing stable and high yield expression of fusion polypeptides of both  
 CC eukaryotic and prokaryotic origin and to encode a protein product for use  
 CC as an antigen for detecting serum antibodies. The presence of serum  
 CC antibodies to M. tuberculosis antigens in an individual indicates that  
 CC the individual is infected with it. The fusion polypeptides are useful as  
 CC sources of proteins for monitoring binding of serum antibodies to fusion  
 CC proteins and as an immunogen to induce and/or enhance immune responses.  
 CC The coding sequences can be ligated with a coding sequence of another  
 CC molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and  
 CC can be used in vivo as a DNA vaccine. This polynucleotide sequence  
 CC represents the DNA encoding the Ral2-DPPD fusion protein.  
 XX  
 SQ Sequence 702 BP; 127 A; 229 C; 225 G; 121 T; 0 other;

Query Match 100.0%; Score 396; DB 23; Length 702;

Best Local Similarity 100.0%; Pred. No. 1.2e-84;

Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGTGGGCGAGGATTCGCCATTCCGATC 60  
 DB 25 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGTGGGCGAGGATTCGCCATTCCGATC 84  
 QY 61 GGGCAGCGATGGCGATCGCGGCGGCGAGATCCGATCGGTGGGCGTCAACCGGCGACGA 120  
 DB 85 GGGCAGCGATGGCGATCGCGGCGGCGAGATCCGATCGGTGGGCGTCAACCGGCGACGA 144  
 QY 121 ATCGGCGCTACCGCTTCTCGGCTGGGTGTTGCGACAAACGCGGCGGCGACGA 180  
 DB 145 ATCGGCGCTACCGCTTCTCGGCTGGGTGTTGCGACAAACGCGGCGGCGACGA 204  
 QY 181 GTCCAAACGGTGGTGGGAGCGCTCGGGGCGGCAAGTCTCGGATCTCCACCGGCGACGFG 240  
 DB 205 GTCCAAACGGTGGTGGGAGCGCTCGGGGCGGCAAGTCTCGGATCTCCACCGGCGACGFG 264  
 QY 241 ATCACCGGCTCGACGGCTCCGATCAACTCGGCCACCGGATGGCGGACGCGGTTAAC 300  
 DB 265 ATCACCGGCTCGACGGCTCCGATCAACTCGGCCACCGGATGGCGGACGCGGTTAAC 324  
 QY 301 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTCCGGCGGCGACGGT 360  
 DB 325 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTCCGGCGGCGACGGT 384  
 QY 361 ACAGGGAACGTGACATTTGGCCGAGGACCCCGGCC 396  
 DB 385 ACAGGGAACGTGACATTTGGCCGAGGACCCCGGCC 420

## RESULT 5

ABK14140  
 ID ABK14140 standard; DNA; 702 BP.

XX AC ABK14140;

XX DT 08-MAY-2002 (first entry)

XX DE DNA encoding antigenic fusion protein Ral2-DPPD (Mtb24).

XX KW Fusion protein; tuberculosis; Mycobacterium tuberculosis; gene; ds;

XX KW tuberculostatic; immunogen; vaccine; Ral2-DPPD; Mtb24.

XX OS Chimeric - Mycobacterium tuberculosis.

XX Key Location/Qualifiers

FT CDS 1..702

FT FT /tag= a

FT FT /product= "Mtb24 #1"

FT FT /partial

FT FT /transl\_except= (pos:694..696, aa:Xaa)

FT FT /note= "No start or stop codon. Xaa= In frame stop codon"

FT CDS 2..702

FT FT /tag= b  
 FT FT /product= "Mtb24 #2"  
 FT FT /partial  
 FT FT /note= "No start or stop codon"  
 FT FT /transl\_except= (pos:263..265, aa:Xaa)  
 FT FT /transl\_except= (pos:353..355, aa:Xaa)  
 FT FT /transl\_except= (pos:395..397, aa:Xaa)  
 FT FT /transl\_except= (pos:470..472, aa:Xaa)  
 FT FT /transl\_except= (pos:701..702, aa:Ser)  
 FT FT /note= "This codon has an apparent 1 nucleotide deletion  
 FT which alters the reading frame."  
 FT Xaa= In frame stop codon"  
 FT CDS 3..701  
 FT FT /tag= c  
 FT FT /product= "Mtb24 #3"  
 FT FT /partial  
 FT FT /note= "No start or stop codon"  
 FT FT /transl\_except= (pos:1..2, aa:Pro)  
 FT FT /transl\_except= (pos:39..41, aa:Xaa)  
 FT FT /transl\_except= (pos:321..323, aa:Xaa)  
 FT FT /transl\_except= (pos:339..341, aa:Xaa)  
 FT FT /transl\_except= (pos:450..452, aa:Xaa)  
 FT FT /transl\_except= (pos:621..623, aa:Xaa)  
 FT FT /note= "No start or stop codon. Xaa= In frame stop codon"  
 XX US2002009459-A1.  
 XX PD 24-JAN-2002.  
 XX PF 07-APR-1999; 99US-0287849.  
 XX PR 13-MAR-1997; 97US-0818112.  
 XX PR 01-OCT-1997; 97US-0942578.  
 XX PR 18-FEB-1998; 98US-0025197.  
 XX PR 07-APR-1998; 98US-0056556.  
 XX PR 30-DEC-1998; 98US-0223040.  
 XX (REED/) REED S G.  
 XX PA (SKEI/) SKEIKY Y A.  
 XX PA (DILL/) DILLON D C.  
 XX PA (ALDE/) ALDERSON M.  
 XX PA (CAMP/) CAMPOS-NETO A.  
 XX PI Reed SG, Skeiky YA, Dillon DC, Alderson M, Campos-Neto A;  
 XX WPI; 2002-171134/22.  
 XX P-PSDB; AAU74600, AAU76541, AAU76542.  
 XX PT New fusion proteins of Mycobacterium tuberculosis antigens, useful for  
 XX diagnosing, treating or preventing M. tuberculosis infection, .  
 XX PT particularly as vaccine for treating or preventing tuberculosis .  
 XX PS Examples: Fig 13; 62pp; English.  
 XX CC The invention relates to a purified polypeptide which induces an immune  
 XX response of Mycobacterium tuberculosis. Polypeptides of the invention are  
 XX useful for diagnosing, treating or preventing M. tuberculosis infection,  
 XX particularly tuberculosis infection. In particular, the polypeptides are  
 XX useful as a vaccine formulation with an adjuvant to afford long-term  
 XX protection in animals against the development of tuberculosis. The  
 XX protein coding sequence may be used to encode a protein product for use  
 XX as an immunogen to induce and/or enhance an immune response to M.  
 XX tuberculosis. This sequence represents DNA encoding an M. tuberculosis  
 XX fusion protein of the invention. This polynucleotide encodes 3 different  
 XX proteins, each in a different reading frame.  
 XX SQ Sequence 702 BP; 127 A; 229 C; 225 G; 121 T; 0 other;

Query Match 100.0%; Score 396; DB 24; Length 702;

Best Local Similarity 100.0%; Pred. No. 1.2e-84;

Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGCGGCGGAGGATTCGCCATTCCGATC 60

```
Db 25 |ACGGCCGGTCCGATAAATTCAGCTGTCCAGAGGTGGCAGGATTCCGCAATCCGATC 84
Qy 61 |GGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGTTGGGGTCAACCACCGTTTCAT 120
Db 85 |GGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGTTGGGGTCAACCACCGTTTCAT 144
Qy 121 |ATCGGGGCTACCGCTTCTCTCGGCTTGGGTGTTGTCGACAAACAGCGGCGCACGA 180
Db 145 |ATCGGGGCTACCGCTTCTCTCGGCTTGGGTGTTGTCGACAAACAGCGGCGCACGA 204
Qy 181 |GTCCAAACGCTGGTTCGGAGGCGCTCCGGCGCAAGTCTCGCATCTCCACCGGCGAGTG 240
Db 205 |GTCCAAACGCTGGTTCGGAGGCGCTCCGGCGCAAGTCTCGCATCTCCACCGGCGAGTG 264
Qy 241 |ATCACCGGCTCGACGGCGCTCCGATCAACTCGGCCACCGGATGGCGGCGCGTTAAAC 300
Db 265 |ATCACCGGCTCGACGGCGCTCCGATCAACTCGGCCACCGGATGGCGGCGCGTTAAAC 324
Qy 301 |GGGCATCATCCGCTGAGCTCATCTCGGTGACCTGGCAAAACAGTTCGGGCGGCGACGGT 360
Db 325 |GGGCATCATCCGCTGAGCTCATCTCGGTGACCTGGCAAAACAGTTCGGGCGGCGACGGT 384
Qy 361 |ACAGGAACGTCGACATTGGCCGAGGAGGACCCCGGCC 396
Db 385 |ACAGGAACGTCGACATTGGCCGAGGAGGACCCCGGCC 420
```

## RESULT 6

```
AAD47077
ID AAD47077 standard; DNA; 1002 BP.
XX
AC
XX
AD AAD47077;
DT 27-JAN-2003 (first entry)
DE Mycobacterium tuberculosis mature Ra35 antigen encoding DNA.
KW Vaccine; immunity; diagnostic agent; gene therapy; Ra35 antigen;
KW gene; ds.
OS Mycobacterium tuberculosis.
FH Key Location/Qualifiers
FT CDS 4..996
FT FT /*tag= a
FT FT /product= "Ra35 mature antigenic protein"
FT FT /transl_except= (pos:547..549, aa:Asp)
FT FT /transl_except= (pos:550..552, aa:Ser)
```

## WO200272792-A2.

```
XX
XX
XX 19-SEP-2002.
XX
XX 13-MAR-2002; 2002WO-US08223.
XX
XX 13-MAR-2001; 2001US-275837P.
XX. (CORI-) CORIXA CORP.
XX
XX Skeiky Y, Brannon M, Guderian J;
XX
XX WPI; 2002-759844/82.
DR P-PSDB; AAE29702.
XX
XX New recombinant nucleic acid molecule comprising a Leishmania TSA,
XX LeIF, M15 or 6H polynucleotide, useful as vaccine to elicit protective
XX immunity against pathogenic microorganisms e.g. Leishmania and
XX Mycobacterium tuberculosis -
XX
XX Disclosure; Page 79; 155pp; English.
XX
XX The invention relates to a recombinant nucleic acid molecule encoding a
```

```
CC fusion polypeptide. The recombinant nucleic acid comprises a heterologous
CC polynucleotide sequence encoding an antigen or an antigenic fragment from
CC Mycobacterium sp. and a Leishmania polynucleotide sequence encoding a
CC polypeptide or its fragment. The Leishmania polynucleotide is selected
CC from TSA, LeIF, M15, and 6H polynucleotides. Sequences of the invention
CC are used in methods for eliciting immune response in mammals. They are
CC useful as vaccines to elicit protective immunity against pathogenic
CC microorganisms such as Leishmania and Mycobacterium tuberculosis. Fusion
CC polypeptides are used for enhancing the expression of polynucleotides,
CC as in vivo diagnostic agents and for raising antibodies in a non-human
CC animal. The invention is used in gene therapy. The present sequence is
CC M. tuberculosis mature Ra35 (N-terminus of MTB32A; RA35FL) antigen
CC encoding DNA.
XX
```

Sequence 1002 BP; 169 A; 331 C; 342 G; 160 T; 0 other;

```
Query Match 100.0%; Score 396; DB 24; Length 1002;
Best Local Similarity 100.0%; Pred. No. 1.2e-84;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ACGGCCGCTCCGATAAATTCAGCTGTCCAGAGGTGGCAGGATTCCGCAATCCGATC 60
Db 598 ACGGCCGCTCCGATAAATTCAGCTGTCCAGAGGTGGCAGGATTCCGCAATCCGATC 657
Qy 61 GGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGTTGGGGTCAACCACCGTTTCAT 120
Db 658 GGCAGGCGATGGCGATCGCGGGCCAGATCGGTTGGGGTCAACCACCGTTTCAT 717
Qy 121 ATCGGGGCTACCGCTTCTCTCGGCTTGGGTGTTGTCGACAAACAGCGGCGCACGA 180
Db 718 ATCGGGGCTACCGCTTCTCTCGGCTTGGGTGTTGTCGACAAACAGCGGCGCACGA 777
Qy 181 GTCCAAACGCTGGTTCGGAGGCGCTCCGGCGCAAGTCTCGCATCTCCACCGGCGAGTG 240
Db 778 GTCCAAACGCTGGTTCGGAGGCGCTCCGGCGCAAGTCTCGCATCTCCACCGGCGAGTG 837
Qy 241 ATCACCGGCTCGACGGCGCTCCGATCAACTCGGCCACCGGATGGCGGCGCGTTAAAC 300
Db 838 ATCACCGGCTCGACGGCGCTCCGATCAACTCGGCCACCGGATGGCGGCGCGTTAAAC 897
Qy 301 GGCATCATCCGCTGAGCTCATCTCGGTGACCTGGCAAAACAGTTCGGGCGGCGACGGT 360
Db 898 GGCATCATCCGCTGAGCTCATCTCGGTGACCTGGCAAAACAGTTCGGGCGGCGACGGT 957
Qy 361 ACAGGAACGTCGACATTGGCCGAGGAGGACCCCGGCC 396
Db 958 ACAGGAACGTCGACATTGGCCGAGGAGGACCCCGGCC 993
```

## RESULT 7

```
AAD47078
ID AAD47078 standard; DNA; 1002 BP.
XX
AC AAD47078;
XX
XX 27-JAN-2003 (first entry)
DE Mycobacterium tuberculosis Ra35FLMutSA mutant antigenic protein DNA.
KW Vaccine; immunity; diagnostic agent; gene therapy; Ra35FLMutSA; mutant;
KW gene; antigen; ds.
OS Mycobacterium tuberculosis.
XX Synthetic.
XX
XX Key Location/Qualifiers
XX CDS 4..996
XX FT /*tag= a
XX FT /product= "Ra35FLMutSA mutant antigenic protein"
XX
XX WO200272792-A2.
XX
XX 19-SEP-2002.
```

```
XX 13-MAR-2002; 2002WO-US08223.
XX
XX 13-MAR-2001; 2001US-275837P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Skelky Y, Brannon M, Guderian J;
XX
XX WPI: 2002-759844/82.
XX P-PSDB; AAE29703.
XX
XX New recombinant nucleic acid molecule comprising a Leishmania TSA,
XX Leif, M15 or 6H polynucleotide, useful as vaccine to elicit protective
XX immunity against pathogenic microorganisms e.g. Leishmania and
XX Mycobacterium tuberculosis
XX
XX Disclosure; Page 80-81; 155pp; English.
XX
XX The invention relates to a recombinant nucleic acid molecule encoding a
XX fusion polypeptide. The recombinant nucleic acid comprises a heterologous
XX polynucleotide sequence encoding an antigen or an antigenic fragment from
XX Mycobacterium sp. and a Leishmania polynucleotide sequence encoding a
XX polypeptide or its fragment. The Leishmania polynucleotide is selected
XX from TSA, Leif, M15, and 6H polynucleotides. Sequences of the invention
XX are used in methods for eliciting immune response in mammals. They are
XX useful as vaccines to elicit protective immunity against pathogenic
XX microorganisms such as Leishmania and Mycobacterium tuberculosis. Fusion
XX polypeptides are used for enhancing the expression of polynucleotides,
XX as in vivo diagnostic agents and for raising antibodies in a non-human
XX animal. The invention is used in gene therapy. The present sequence is
XX M. tuberculosis Ra35FLmutSA mutant antigenic protein encoding DNA.
XX
XX Sequence 1002 BP; 169 A; 331 C; 341 G; 161 T; 0 other;
XX
XX Query Match 100.0%; Score 396; DB 24; Length 1002;
XX Best Local Similarity 100.0%; Pred. No. 1.2e-84;
XX Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 ACGGGCGGCTCGGATAACTTCCAGCTGTCCAGGGTGGCAGGGATTCCCGATTCGGATC 60
DB 598 ACGGGCGGCTCGGATAACTTCCAGCTGTCCAGGGTGGCAGGGATTCCCGATTCGGATC 657
QY 61 GGGCAGGCATGGCGATCGCGGGCCAGATCCGATCGGTGGGGGTCCACCCACCGTTTCA 120
DB 658 GGGCAGGCATGGCGATCGCGGGCCAGATCCGATCGGTGGGGGTCCACCCACCGTTTCA 717
QY 121 ATCGGGCCCTACCGCCCTTCTCGGCTTGGGTGTCGACAAACGCGGCAACGCGCAGCA 180
DB 718 ATCGGGCCCTACCGCCCTTCTCGGCTTGGGTGTCGACAAACGCGGCAACGCGCAGCA 777
QY 181 GTCCAAACGGTGGTGGGAGCGCTCCGGCGGCAAGTCTGGCATCTCCACCGCGGACGTG 240
DB 778 GTCCAAACGGTGGTGGGAGCGCTCCGGCGGCAAGTCTGGCATCTCCACCGCGGACGTG 837
QY 241 ATCAGCGGGTGGTGGGAGCGCTCCGGTCACTCGGCCACCGCATCGGCGGACGCGTTTAA 300
DB 838 ATCAGCGGGTGGTGGGAGCGCTCCGGTCACTCGGCCACCGCATCGGCGGACGCGTTTAA 897
QY 301 GGGCATCATCCGGTGGTGGGAGCGCTCGGTGACCTGGCAACCAAGTCCGGCGGCGCAGCG 360
DB 898 GGGCATCATCCGGTGGTGGGAGCGCTCGGTGACCTGGCAACCAAGTCCGGCGGCGCAGCG 957
QY 361 ACAGGGAAGCTGACATTTGGCGGAGGACCCCGGCC 396
DB 958 ACAGGGAAGCTGACATTTGGCGGAGGACCCCGGCC 993
XX
RESULT 8
AAD28336
ID AAD28336 standard; cDNA; 1002 BP.
XX
XX AAD28336;
```

```
XX 22-APR-2002 (first entry)
XX
XX Mycobacterium sp. MTB32A (Ra35FL) mature protein encoding cDNA.
XX
XX Fusion protein; antigen; serological sensitivity; immune response;
XX tuberculosis; infection; vaccine; MTB32A; Ra32FL; ss.
XX
XX Mycobacterium sp.
XX
XX Key Location/Qualifiers
XX CDS 4..996
XX
XX /tag- a
XX /product- "Ra35 mature protein"
XX /transl_except- (pos:547..549, aa:Asp)
XX /transl_except- (pos:550..552, aa:Ser)
XX
XX WO200198460-A2.
XX
XX 27-DEC-2001.
XX
XX 20-JUN-2001; 2001WO-US19959.
XX
XX 20-JUN-2000; 2000US-0597796.
XX 01-FEB-2001; 2001US-265737P.
XX (CORI-) CORIXA CORP.
XX
XX Skeiky Y, Reed S, Alderson M;
XX
XX WPI: 2002-147798/19.
XX P-PSDB; AAE17566.
XX
XX Composition comprising MTB39 antigen and MTB32A antigen from
XX Mycobacterium species, useful for eliciting immune response in a
XX subject
XX
XX Disclosure; Page 95; 136pp; English.
XX
XX The present invention relates to fusion proteins containing at least
XX two Mycobacterium species antigens, nucleotides encoding them and
XX compositions comprising such fusion proteins. The present invention
XX particularly relates to nucleic acids encoding fusion proteins that
XX include two or more individual M. tuberculosis antigens which increase
XX the serological sensitivity of sera from individuals infected with
XX tuberculosis and methods for their use in diagnosis, prevention and
XX treatment of tuberculosis infection. Sequences of the invention are
XX useful for eliciting an immune response in a mammal, e.g., human,
XX immunised with BCG. They are useful in the diagnosis, treatment and
XX prevention of Mycobacterium infection. The fusion proteins and the
XX polynucleotides are useful as diagnostic tools in patients infected
XX with Mycobacterium, in vitro and in vivo assays for detecting humoral
XX antibodies or cell-mediated immunity against M. tuberculosis, for the
XX diagnosis of an infection or monitoring of disease progression, as
XX immunogens to generate or elicit a protective immune response in a
XX patient and for raising anti-M. tuberculosis antibodies in a non-human
XX animal. Sequences of the invention are also used as vaccines. MTB32A
XX fusion proteins of the invention are useful as in vivo diagnostic agents
XX for intradermal skin test. The present sequence is a cDNA encoding
XX Mycobacterium species MTB32A (Ra32FL) mature protein.
XX
XX Sequence 1002 BP; 169 A; 331 C; 342 G; 160 T; 0 other;
XX
XX Query Match 100.0%; Score 396; DB 24; Length 1002;
XX Best Local Similarity 100.0%; Pred. No. 1.2e-84;
XX Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 ACGGGCGGCTCGGATAACTTCCAGCTGTCCAGGGTGGCAGGGATTCCCGATTCGGATC 60
DB 598 ACGGGCGGCTCGGATAACTTCCAGCTGTCCAGGGTGGCAGGGATTCCCGATTCGGATC 657
QY 61 GGGCAGGCATGGCGATCGCGGGCCAGATCCGATCGGTGGGGGTCCACCCACCGTTTCA 120
DB 658 GGGCAGGCATGGCGATCGCGGGCCAGATCCGATCGGTGGGGGTCCACCCACCGTTTCA 717
QY 121 ATCGGGCCCTACCGCCCTTCTCGGCTTGGGTGTCGACAAACGCGGCAACGCGCAGCA 180
DB 718 ATCGGGCCCTACCGCCCTTCTCGGCTTGGGTGTCGACAAACGCGGCAACGCGCAGCA 777
QY 181 GTCCAAACGGTGGTGGGAGCGCTCCGGCGGCAAGTCTGGCATCTCCACCGCGGACGTG 240
DB 778 GTCCAAACGGTGGTGGGAGCGCTCCGGCGGCAAGTCTGGCATCTCCACCGCGGACGTG 837
QY 241 ATCAGCGGGTGGTGGGAGCGCTCCGGTCACTCGGCCACCGCATCGGCGGACGCGTTTAA 300
DB 838 ATCAGCGGGTGGTGGGAGCGCTCCGGTCACTCGGCCACCGCATCGGCGGACGCGTTTAA 897
QY 301 GGGCATCATCCGGTGGTGGGAGCGCTCGGTGACCTGGCAACCAAGTCCGGCGGCGCAGCG 360
DB 898 GGGCATCATCCGGTGGTGGGAGCGCTCGGTGACCTGGCAACCAAGTCCGGCGGCGCAGCG 957
QY 361 ACAGGGAAGCTGACATTTGGCGGAGGACCCCGGCC 396
DB 958 ACAGGGAAGCTGACATTTGGCGGAGGACCCCGGCC 993
XX
RESULT 8
AAD28336
ID AAD28336 standard; cDNA; 1002 BP.
XX
XX AAD28336;
```

```
Db 658 GGGCAGCGATGGCGATCGCGGCCAGATCCGATCGGGTGGGGGTCAACCCACCGTTTCAT 717
Qy 121 ATCGGGGCTACCGCTTCCTCGGCTTGGGTGTTGTCACAAACGCAACGGCGACGA 180
Db 718 ATCGGGGCTACCGCTTCCTCGGCTTGGGTGTTGTCACAAACGCAACGGCGACGA 777
Qy 181 GTCCACGCGTGTGCGGAGCGCTCGCGGCGCAAGTCTCGGCATCTCCACCGCGAGGTG 240
Db 778 GTCCACGCGTGTGCGGAGCGCTCGCGGCGCAAGTCTCGGCATCTCCACCGCGAGGTG 837
Qy 241 ATCACCGCGTGTGCGGAGCGCTCGCATCAACTCGGCCACCGCATGGCGGACGCTTAAC 300
Db 838 ATCACCGCGTGTGCGGAGCGCTCGCATCAACTCGGCCACCGCATGGCGGACGCTTAAC 897
Qy 301 GGGCATCATCCCGTGCAGCTCATCTCGGTGACCTGCGCAACCAAGTCGGCGGCGACGCT 360
Db 898 GGGCATCATCCCGTGCAGCTCATCTCGGTGACCTGCGCAACCAAGTCGGCGGCGACGCT 957
Qy 361 ACAGGNAAGTGACATTGGCCGAGGACCCCGGCC 396
Db 958 ACAGGNAAGTGACATTGGCCGAGGACCCCGGCC 993

RESULT 9
AAD28337
ID AAD28337 standard; cDNA; 1002 BP.
XX
AC AAD28337;
XX
DT 22-APR-2002 (first entry)
DE
KW Mycobacterium species Ra35FLMutSA mutant cDNA.
XX
KW Fusion protein; antigen; serological sensitivity; immune response;
KW tuberculosis; infection; vaccine; Ra32FLMutSA; mutant; mutein; ss.
XX
OS Mycobacterium sp.
XX
FH Key Location/Qualifiers
FT CDS 4..996
FT FT /*tag= a
FT FT /product= "Ra35FLMutSA protein"
FT FT
XX
PN W0200198460-A2.
XX
XX 27-DEC-2001.
XX
XX 20-JUN-2001; 2001WO-US19959.
XX
PR 20-JUN-2000; 2000US-0597796.
PR 01-FEB-2001; 2001US-265737P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky Y, Reed S, Alderson M;
XX
XX WPI; 2002-147798/19.
DR P-PSDB; AAEL7567.
XX
XX
PT Composition comprising MTB39 antigen and MTB32A antigen from
PT Mycobacterium species, useful for eliciting immune response in a
PT subject -
XX
XX
PS Claim 67; Page 96-97; 136pp; English.
XX
CC The present invention relates to fusion proteins containing at least
CC two Mycobacterium species antigens, nucleotides encoding them and
CC compositions comprising such fusion proteins. The present invention
CC particularly relates to nucleic acids encoding fusion proteins that
CC include two or more individual M. tuberculosis antigens which increase
CC the serological sensitivity of sera from individuals infected with
CC tuberculosis and methods for their use in diagnosis, prevention and
CC treatment of tuberculosis infection. Sequences of the invention are
```

```
CC useful for eliciting an immune response in a mammal, e.g., human,
CC immunised with BCG. They are useful in the diagnosis, treatment and
CC prevention of Mycobacterium infection. The fusion proteins and the
CC polynucleotides are useful as diagnostic tools in patients infected
CC with Mycobacterium, in vitro and in vivo assays for detecting humoral
CC antibodies or cell-mediated immunity against M. tuberculosis, for the
CC diagnosis of an infection or monitoring of disease progression, as
CC immunogens to generate or elicit a protective immune response in a
CC patient and for raising anti-M. tuberculosis antibodies in a non-human
CC animal. Sequences of the invention are also used as vaccines. MTB32A
CC fusion proteins of the invention are useful as in vivo diagnostic agents
CC for intradermal skin test. The present sequence is a cDNA encoding
CC Mycobacterium species Ra32FL mature protein mutant, Ra32FLMutSA.
XX
SQ Sequence 1002 BP; 169 A; 331 C; 341 G; 161 T; 0 other;
Query Match 100.0%; Score 396; DB 24; Length 1002;
Best Local Similarity 100.0%; Pred. NO. 1.2e-84;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ACGGCCGCGTCCGATTAACCTTCCAGCTGTCCAGGGTGGGCGAGGATTCGCCATTCCCGATC 60
Db 598 ACGGCCGCGTCCGATTAACCTTCCAGCTGTCCAGGGTGGGCGAGGATTCGCCATTCCCGATC 657
Qy 61 GGGCAGCGATGGCGATCGCGGCGCAGATCCGATCGGGTGGGGGTGACCCACCGTTTCAT 120
Db 658 GGGCAGCGATGGCGATCGCGGCGCAGATCCGATCGGGTGGGGGTGACCCACCGTTTCAT 717
Qy 121 ATCGGGCCTACCGCTTCCTCGGCTTGGGTGTTGTCACAAACGCAACGGCGACGA 180
Db 718 ATCGGGCCTACCGCTTCCTCGGCTTGGGTGTTGTCACAAACGCAACGGCGACGA 777
Qy 181 GTCCACGCGTGTGCGGAGCGCTCGCGGCGCAAGTCTCGGCATCTCCACCGCGAGGTG 240
Db 778 GTCCACGCGTGTGCGGAGCGCTCGCGGCGCAAGTCTCGGCATCTCCACCGCGAGGTG 837
Qy 241 ATCACCGCGTGTGCGGAGCGCTCGCATCAACTCGGCCACCGCATGGCGGACGCGCTTAAC 300
Db 838 ATCACCGCGTGTGCGGAGCGCTCGCATCAACTCGGCCACCGCATGGCGGACGCGCTTAAC 897
Qy 301 GGGCATCATCCCGTGCAGCTCATCTCGGTGACCTGCGCAACCAAGTCGGCGGCGACGCT 360
Db 898 GGGCATCATCCCGTGCAGCTCATCTCGGTGACCTGCGCAACCAAGTCGGCGGCGACGCT 957
Qy 361 ACAGGNAAGTGACATTGGCCGAGGACCCCGGCC 396
Db 958 ACAGGNAAGTGACATTGGCCGAGGACCCCGGCC 993

RESULT 10
AA34251
ID AA34251 standard; DNA; 1068 BP.
XX
AC AA34251;
XX
XX 06-JUL-1999 (first entry)
XX
DE Mycobacterium species nucleic acid sequence 50D.
XX
KW Secreted protein; Mycobacterium; primer; PCR; amplification; probe;
KW hybridisation; detection; vaccine; immunisation; infection; ss.
XX
OS Mycobacterium sp.
XX
XX WO9909186-A2.
XX
XX 25-FEB-1999.
XX
XX 14-AUG-1998; 98WO-FR01813.
PR 11-SEP-1997; 97FR-0011325.
PR 14-AUG-1997; 97FR-0010404.
XX
```

```
PA (INSP ) INST PASTEUR.
XX
PI Gicquel B, Lim EM, Pellicic V, Portnoi D, Goguet de la Salmoniere Y;
PI Guigueno A;
XX
PR 11-SEP-1997; 97FR-0011325.
PR 14-AUG-1997; 97FR-0010404.
XX
PA (INSP ) INST PASTEUR.
XX
PI Gicquel B, Lim EM, Pellicic V, Portnoi D, Goguet de la Salmoniere Y;
PI Guigueno A;
XX
DR WPI: 1999-181045/15.
DR P-PSDB; AAY05000.
XX
XX Mycobacterial DNA vectors containing reporter constructs - for
PT identifying coding or promoter sequences involved in
PT infection-associated protein expression
XX
PS Claim 22; Fig 50P; 309pp; French.
XX
XX Sequences AAX34001-X34252 represent nucleic acids encoding secreted
CC proteins from various Mycobacterium species microorganisms. The
CC nucleotide sequences can be used as primers and probes for methods
CC for detecting and identifying mycobacteria, especially belonging to
CC the M. tuberculosis complex. The encoded proteins can be used in
CC vaccines for immunisation against a bacterial or viral infection.
XX
XX - Sequence 1068 BP; 169 A; 352 C; 375 G; 172 T; 0 other;
XX
Query Match 100.0%; Score 396; DB 20; Length 1068;
Best Local Similarity 100.0%; Pred. No. 1.2e-84;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACGGCGCGGTCCGATAACTTCCAGCTGTCACAGGTGGGCGAGGATTCGCCATTCGGATC 60
DB 670 ACGGCGCGGTCCGATAACTTCCAGCTGTCACAGGTGGGCGAGGATTCGCCATTCGGATC 729
QY 61 GGGCAGGCGATGGCGATCGGGGCCAGATCCGATCGGTGGGGGTACCCACCGTTTCAT 120
DB 730 GGGCAGGCGATGGCGATCGGGGCCAGATCCGATCGGTGGGGGTACCCACCGTTTCAT 789
QY 121 ATCGGGCTACGGCTTCTCGGCTTGGGTGTCGACAAACACGCAACGGCGCACGA 180
DB 790 ATCGGGCTACGGCTTCTCGGCTTGGGTGTCGACAAACACGCAACGGCGCACGA 849
QY 181 GTCCAAACCGTGGTGGGAGGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
DB 850 GTCCAAACCGTGGTGGGAGGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGGCGACGTG 909
QY 241 ATCACCAGGCGTGGCGGCGTCCGATCACTCGGCGCCAGCGGATGGCGGACGGTTAAC 300
DB 910 ATCACCAGGCGTGGCGGCGTCCGATCACTCGGCGCCAGCGGATGGCGGACGGTTAAC 969
QY 301 GGGCATATCCGGTGACGTATCTCGTGACCTGGCAACCAAGTCGGGCGGCGACGCGT 360
DB 970 GGGCATATCCGGTGACGTATCTCGTGACCTGGCAACCAAGTCGGGCGGCGACGCGT 1029
QY 361 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 396
DB 1030 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 1065
RESULT 11
AAX34252
ID AAX34252 standard; DNA; 1143 BP.
XX
AC AAX34252;
XX
DT 06-JUL-1999 (first entry)
XX
DE Mycobacterium species nucleic acid sequence 50F.
XX
KW Secreted protein; Mycobacterium; primer; PCR; amplification; probe;
KW hybridisation; detection; vaccine; immunisation; infection; ss.
XX
OS Mycobacterium sp.
XX
PN W09909186-A2.
XX
```

---

```
PD 25-FEB-1999.
XX
PF 14-AUG-1998; 98WO-FR01813.
XX
PR 11-SEP-1997; 97FR-0011325.
PR 14-AUG-1997; 97FR-0010404.
XX
PA (INSP ) INST PASTEUR.
XX
PI Gicquel B, Lim EM, Pellicic V, Portnoi D, Goguet de la Salmoniere Y;
PI Guigueno A;
XX
DR WPI: 1999-181045/15.
DR P-PSDB; AAY04830.
XX
XX Mycobacterial DNA vectors containing reporter constructs - for
PT identifying coding or promoter sequences involved in
PT infection-associated protein expression
XX
PS Claim 22; Fig 50P; 309pp; French.
XX
XX Sequences AAX34001-X34252 represent nucleic acids encoding secreted
CC proteins from various Mycobacterium species microorganisms. The
CC nucleotide sequences can be used as primers and probes for methods
CC for detecting and identifying mycobacteria, especially belonging to
CC the M. tuberculosis complex. The encoded proteins can be used in
CC vaccines for immunisation against a bacterial or viral infection.
XX
XX - Sequence 1143 BP; 189 A; 373 C; 395 G; 186 T; 0 other;
XX
Query Match 100.0%; Score 396; DB 20; Length 1143;
Best Local Similarity 100.0%; Pred. No. 1.2e-84;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACGGCGCGGTCCGATAACTTCCAGCTGTCACAGGTGGGCGAGGATTCGCCATTCGGATC 60
DB 745 ACGGCGCGGTCCGATAACTTCCAGCTGTCACAGGTGGGCGAGGATTCGCCATTCGGATC 804
QY 61 GGGCAGGCGATGGCGATCGGGGCCAGATCCGATCGGTGGGGGTACCCACCGTTTCAT 120
DB 805 GGGCAGGCGATGGCGATCGGGGCCAGATCCGATCGGTGGGGGTACCCACCGTTTCAT 864
QY 121 ATCGGGCTACGGCTTCTCGGCTTGGGTGTCGACAAACACGCAACGGCGCACGA 180
DB 865 ATCGGGCTACGGCTTCTCGGCTTGGGTGTCGACAAACACGCAACGGCGCACGA 924
QY 181 GTCCAAACCGTGGTGGGAGGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
DB 925 GTCCAAACCGTGGTGGGAGGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGGCGACGTG 984
QY 241 ATCACCAGGCGTGGCGGCGTCCGATCACTCGGCGCCAGCGGATGGCGGACGGTTAAC 300
DB 985 ATCACCAGGCGTGGCGGCGTCCGATCACTCGGCGCCAGCGGATGGCGGACGGTTAAC 1044
QY 301 GGGCATATCCGGTGACGTATCTCGTGACCTGGCAACCAAGTCGGGCGGCGACGCGT 360
DB 1045 GGGCATATCCGGTGACGTATCTCGTGACCTGGCAACCAAGTCGGGCGGCGACGCGT 1104
QY 361 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 396
DB 1105 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 1140
RESULT 12
AAL40771
ID AAL40771 standard; DNA; 1742 BP.
XX
AC AAL40771;
XX
DT 03-OCT-2002 (first entry)
XX
DE Nucleotide sequence encoding Ral2-WT1 fusion protein.
XX
```





```
Query Match      100.0%; Score 396; DB 24; Length 2190;
Best Local Similarity 100.0%; Pred. No. 1.3e-84;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGGCGCGTCCGATAAATTCAGCTGTCCAGGTGGCGAGGATTCGCCATTCGGATC 60
DB 22 ACGGCGCGTCCGATAAATTCAGCTGTCCAGGTGGCGAGGATTCGCCATTCGGATC 81
QY 61 GGGCAGCGATGGCGATCGCGGCGCAGATCCGATCGGTGGGGGTTCACCCACCGTTTAT 120
DB 82 GGGCAGCGATGGCGATCGCGGCGCAGATCCGATCGGTGGGGGTTCACCCACCGTTTAT 141
QY 121 ATCGGCGCTACCGCTTCCTCGGCTTGGGTGTGTCGACAAACGCGACGCGACGA 180
DB 142 ATCGGCGCTACCGCTTCCTCGGCTTGGGTGTGTCGACAAACGCGACGCGACGA 201
QY 181 GTCCAAACGGTGTGGGAGCGCTCGCGGCGCAAGTCTCGGATCTCCACCGCGACGNG 240
DB 202 GTCCAAACGGTGTGGGAGCGCTCGCGGCGCAAGTCTCGGATCTCCACCGCGACGNG 261
QY 241 ATCACCGCGTTCGACGGCGCTCCGATCAACTCGGCAACCGCGATCGCGACGCGTTAAC 300
DB 262 ATCACCGCGTTCGACGGCGCTCCGATCAACTCGGCAACCGCGATCGCGACGCGTTAAC 321
QY 301 GGGCATCATCCCGGTGAGGTCTATCTCGGTGACCTCGGCAACCAAGTCGGGCGGACGCGT 360
DB 322 GGGCATCATCCCGGTGAGGTCTATCTCGGTGACCTCGGCAACCAAGTCGGGCGGACGCGT 381
QY 361 ACAGGGAACGTGACATTTGGCGGAGGACCCCGGCC 396
DB 382 ACAGGGAACGTGACATTTGGCGGAGGACCCCGGCC 417
```

## RESULT 14

```
AAD28343
ID AAD28343 standard; DNA; 2190 BP.
XX
AC AAD28343;
XX
XX
XX 22-APR-2002 (first entry)
DT
XX
DE Mycobacterium species MTB72FmutSA fusion protein encoding DNA.
DE
KW Fusion protein; antigen; serological sensitivity; immune response;
KW tuberculosis; infection; vaccine; MTB72FmutSA; Ra12-TbH9-Ra35mutSA;
KW mutant; mutin; ds.
XX
OS Mycobacterium sp.
XX
XX
XX Key Location/Qualifiers
FH CDS 1..2190
FT /*tag= a
FT /product= "MTB72FmutSA fusion protein"
FT 22..417
FT misc_feature /*tag= b
FT /*note= "Ra12 DNA fragment"
FT 424..1596
FT misc_feature /*tag= c
FT /*note= "TbH9FL DNA fragment"
FT 1603..2187
FT misc_feature /*tag= d
FT /*note= "Ra35 DNA fragment"
FT mutation replace (2128, T)
FT /*tag= e
XX
XX WO200198460-A2.
XX
XX 27-DEC-2001.
PD
XX
XX 20-JUN-2001; 2001WO-US19959.
XX
XX 20-JUN-2000; 2000US-0597796.
XX
XX 01-FEB-2001; 2001US-265737P.
PR
```

```
XX (CORI-) CORIXA CORP.
PA Skeiky Y, Reed S, Alderson M;
XX WPI: 2002-147798/19.
DR P-PSDB; AAE17573.
XX
XX Composition comprising MTB39 antigen and MTB32A antigen from
PT Mycobacterium species, useful for eliciting immune response in a
PT subject
XX
XX Claim 81; Page 108-109; 136pp; English.
XX
XX The present invention relates to fusion proteins containing at least
CC two Mycobacterium species antigens, nucleotides encoding them and
CC compositions comprising such fusion proteins. The present invention
CC particularly relates to nucleic acids encoding fusion proteins that
CC include two or more individual M. tuberculosis antigens which increase
CC the serological sensitivity of sera from individuals infected with
CC tuberculosis and methods for their use in diagnosis, prevention and
CC treatment of tuberculosis infection. Sequences of the invention are
CC useful for eliciting an immune response in a mammal, e.g., human,
CC immunised with BCG. They are useful in the diagnosis, treatment and
CC prevention of Mycobacterium infection. The fusion proteins and the
CC polynucleotides are useful as diagnostic tools in patients infected
CC with Mycobacterium, in vitro and in vivo assays for detecting humoral
CC antibodies or cell-mediated immunity against M. tuberculosis, for the
CC diagnosis of an infection or monitoring of disease progression, as
CC immunogens to generate or elicit a protective immune response in a
CC patient and for raising anti-M. tuberculosis antibodies in a non-human
CC animal. Sequences of the invention are also used as vaccines. MTB32A
CC fusion proteins of the invention are useful as in vivo diagnostic agents
CC for intradermal skin test. The present sequence is a DNA encoding
CC Mycobacterium species MTB72FmutSA (Ra12-TbH9-Ra35mutSA) mutant fusion
CC protein.
XX
XX Sequence 2190 BP; 340 A; 697 C; 797 G; 356 T; 0 other;
```

```
Query Match      100.0%; Score 396; DB 24; Length 2190;
Best Local Similarity 100.0%; Pred. No. 1.3e-84;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGGCGCGTCCGATAAATTCAGCTGTCCAGGTGGCGAGGATTCGCCATTCGGATC 60
DB 22 ACGGCGCGTCCGATAAATTCAGCTGTCCAGGTGGCGAGGATTCGCCATTCGGATC 81
QY 61 GGGCAGCGATGGCGATCGCGGCGCAGATCCGATCGGTGGGGGTTCACCCACCGTTTAT 120
DB 82 GGGCAGCGATGGCGATCGCGGCGCAGATCCGATCGGTGGGGGTTCACCCACCGTTTAT 141
QY 121 ATCGGCGCTACCGCTTCCTCGGCTTGGGTGTGTCGACAAACGCGACGCGACGA 180
DB 142 ATCGGCGCTACCGCTTCCTCGGCTTGGGTGTGTCGACAAACGCGACGCGACGA 201
QY 181 GTCCAAACGGTGTGGGAGCGCTCCGCGGCGCAAGTCTCGGATCTCCACCGCGACGNG 240
DB 202 GTCCAAACGGTGTGGGAGCGCTCCGCGGCGCAAGTCTCGGATCTCCACCGCGACGNG 261
QY 241 ATCACCGCGTTCGACGGCGCTCCGATCAACTCGGCAACCGCGATCGCGACGCGTTAAC 300
DB 262 ATCACCGCGTTCGACGGCGCTCCGATCAACTCGGCAACCGCGATCGCGACGCGTTAAC 321
QY 301 GGGCATCATCCCGGTGAGGTCTATCTCGGTGACCTCGGCAACCAAGTCGGGCGGACGCGT 360
DB 322 GGGCATCATCCCGGTGAGGTCTATCTCGGTGACCTCGGCAACCAAGTCGGGCGGACGCGT 381
QY 361 ACAGGGAACGTGACATTTGGCGGAGGACCCCGGCC 396
DB 382 ACAGGGAACGTGACATTTGGCGGAGGACCCCGGCC 417
```

RESULT 15

AA140773  
ID AAL40773 standard; DNA; 2191 BP.

AC AAL40773;

DT 03-OCT-2002 (first entry)

DE Nucleotide sequence encoding Ral2-H9-32A fusion protein.

XX Ral2; serine protease antigen; MTB32A; Mycobacterium tuberculosis;  
KW vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;  
KW immunogen; cytokine; gene; ds.

XX Chimeric - Unidentified.

XX Key Location/Qualifiers

FT 1..2190

FT /tag= a

FT /product= "Ral2-H9-32A fusion protein"

XX WO200125401-A2;

XX 12-APR-2001.

XX 06-OCT-2000; 2000WO-US27652.

XX 07-OCT-1999; 99US-0158585.

XX (CORI-) CORIXA CORP.

XX Skeiky Y, Guderian J;

XX WPI; 2001-266299/27.

XX P-PSDB; AAO22142.

XX Recombinant nucleic acid molecule for producing high yield expression  
PT of desired fusion polypeptides, encodes fusion polypeptide comprising  
PT Mycobacterium tuberculosis coding sequence and heterologous polypeptide

XX Disclosure; Fig 6; 39pp; English.

XX The invention relates to a recombinant nucleic acid molecule encoding a  
CC fusion polypeptide, comprising a polynucleotide sequence of Ral2, a 14  
CC kDa C-terminal fragment of serine protease antigen MTB32A of  
CC Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.  
CC The recombinant fusion nucleic acids and polypeptides are useful for  
CC providing stable and high yield expression of fusion polypeptides of both  
CC eukaryotic and prokaryotic origin and to encode a protein product for use  
CC as an antigen for detecting serum antibodies. The presence of serum  
CC antibodies to M. tuberculosis antigens in an individual indicates that  
CC the individual is infected with it. The fusion polypeptides are useful as  
CC sources of proteins for monitoring binding of serum antibodies to fusion  
CC proteins and as an immunogen to induce and/or enhance immune responses.  
CC The coding sequences can be ligated with a coding sequence of another  
CC molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and  
CC can be used in vivo as a DNA vaccine. This polynucleotide sequence  
CC represents the DNA encoding the Ral2-H9-32A fusion protein.

XX Sequence 2191 BP; 340 A; 697 C; 797 G; 356 T; 1 other;

Query Match 100.0%; Score 396; DB 23; Length 2191;

Best Local Similarity 100.0%; Pred. NO. 1.3e-84;

Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGGCGCGTCCGATACCTCCAGCTCCCGGGTGGGCGAGGATTCGCCATCCGATC 60

DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
22 ACGGCGCGTCCGATACCTCCAGCTCCCGGGTGGGCGAGGATTCGCCATCCGATC 81

QY 61 GGGCAGGCGATGCGGATCGGGGCGAGATCCGATCGGGTGGGCGGTCACCCACCGTTTCAT 120

DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
82 GGGCAGGCGATGCGGATCGGGGCGAGATCCGATCGGGTGGGCGGTCACCCACCGTTTCAT 141

QY 121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTGACAAACAGGCAACGGCGACGCA 180  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
142 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTGACAAACAGGCAACGGCGACGCA 201  
QY 181 GTCCAAACGGGTGGTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGAGCTG 240  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
202 GTCCAAACGGGTGGTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGAGCTG 261  
QY 241 ATCACCAGCGGTGACGCGCTCCGATCAACTCGGCCACCGCGATGGCGACGGCTTAAC 300  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
262 ATCACCAGCGGTGACGCGCTCCGATCAACTCGGCCACCGCGATGGCGACGGCTTAAC 321  
QY 301 GGGCATCATCCCCGGTGACGTCATCTCGGTGACCTGGCAACCAAGTCGGGCGGCGACGCGT 360  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
322 GGGCATCATCCCCGGTGACGTCATCTCGGTGACCTGGCAACCAAGTCGGGCGGCGACGCGT 381  
QY 361 ACAGGGAACGTGACATTTGGCCGAGGAGGACCCCGGCC 396  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
382 ACAGGGAACGTGACATTTGGCCGAGGAGGACCCCGGCC 417

Search completed: September 5, 2003, 07:24:26  
Job time : 251 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2003, 07:20:04 ; Search time 75 Seconds  
(without alignments)  
2330.503 Million cell updates/sec

Title: US-09-684-215A-3

Perfect score: 396

Sequence: 1 acggccgcgtccgataactt.....tggccgaggacccccggcc 396

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_NA.\*

- 1: /cgn2\_6/ptodata/1/ina/5A-COMB.seq.\*
- 2: /cgn2\_6/ptodata/1/ina/5B-COMB.seq.\*
- 3: /cgn2\_6/ptodata/1/ina/6A-COMB.seq.\*
- 4: /cgn2\_6/ptodata/1/ina/6B-COMB.seq.\*
- 5: /cgn2\_6/ptodata/1/ina/PCTUS-COMB.seq.\*
- 6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	396	100.0	2287	4	US-09-223-040-1
2	396	100.0	4403765	3	US-09-103-840A-2
3	396	100.0	4411529	3	US-09-103-840A-1
4	394.4	99.6	447	3	US-08-818-112-4
5	394.4	99.6	447	4	US-08-818-111-4
6	394.4	99.6	447	4	US-09-056-556-4
7	394.4	99.6	447	4	US-09-072-596-4
8	391.2	98.8	1872	3	US-08-818-112-17
9	391.2	98.8	1872	4	US-08-818-111-17
10	391.2	98.8	1872	4	US-09-056-556-17
11	391.2	98.8	1872	4	US-09-072-596-17
12	354	89.4	822	4	US-09-736-457-1862
13	354	89.4	900	4	US-09-643-597-353
14	354	89.4	900	4	US-09-606-421B-353
15	354	89.4	945	4	US-09-736-457-1861
16	354	89.4	1012	4	US-09-643-597-351
17	354	89.4	1012	4	US-09-606-421B-351
18	354	89.4	1464	4	US-09-620-412C-348
19	354	89.4	1464	4	US-09-598-419-348
20	354	89.4	1557	4	US-09-620-412C-332
21	354	89.4	1557	4	US-09-598-419-332
22	354	89.4	1578	4	US-09-556-877-188
23	354	89.4	1578	4	US-09-620-412C-188
24	354	89.4	1578	4	US-09-598-419-188
25	354	89.4	1752	4	US-09-620-412C-352
26	354	89.4	1752	4	US-09-598-419-352
27	354	89.4	1752	4	US-09-620-412C-336

Sequence 336, App  
Sequence 308, App  
Sequence 308, App  
Sequence 324, App  
Sequence 324, App  
Sequence 316, App  
Sequence 316, App  
Sequence 340, App  
Sequence 340, App  
Sequence 356, App  
Sequence 356, App  
Sequence 312, App  
Sequence 312, App  
Sequence 344, App  
Sequence 344, App  
Sequence 320, App  
Sequence 320, App  
Sequence 320, App  
Sequence 320, App

28 354 89.4 1758 4 US-09-598-419-336  
29 354 89.4 1860 4 US-09-620-412C-308  
30 354 89.4 1860 4 US-09-598-419-308  
31 354 89.4 1896 4 US-09-620-412C-324  
32 354 89.4 1896 4 US-09-598-419-324  
33 354 89.4 1941 4 US-09-620-412C-316  
34 354 89.4 1941 4 US-09-598-419-316  
35 354 89.4 1965 4 US-09-620-412C-340  
36 354 89.4 1965 4 US-09-598-419-340  
37 354 89.4 2052 4 US-09-620-412C-356  
38 354 89.4 2052 4 US-09-598-419-356  
39 354 89.4 2076 4 US-09-620-412C-312  
40 354 89.4 2076 4 US-09-598-419-312  
41 354 89.4 2103 4 US-09-620-412C-344  
42 354 89.4 2103 4 US-09-598-419-344  
43 354 89.4 2148 4 US-09-620-412C-320  
44 354 89.4 2148 4 US-09-620-412C-328  
45 354 89.4 2148 4 US-09-598-419-320

#### ALIGNMENTS

#### RESULT 1

US-09-223-040-1

; Sequence 1, Application US/09223040

; Patent No. 6544522

; GENERAL INFORMATION:

; APPLICANT: Skeiky, Yasir

; APPLICANT: Alderson, Mark

; APPLICANT: Campos-Neto, Antonio

; APPLICANT: Corixa Corporation

; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens

; FILE REFERENCE: 014058-00901005

; CURRENT APPLICATION NUMBER: US/09/223,040

; CURRENT FILING DATE: 1998-12-30

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 1

; LENGTH: 2287

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: tri-fusion

; OTHER INFORMATION: protein Ra12-TBH9-Ra35

; NAME/KEY: modified\_base

; LOCATION: (30)

; OTHER INFORMATION: n = g, a, c o r t

; NAME/KEY: modified\_base

; LOCATION: (33)

; OTHER INFORMATION: n = g, a, c o r t

; NAME/KEY: CDS

; LOCATION: (42)..(2231)

; NAME/KEY: modified\_base

; LOCATION: (2270)

; OTHER INFORMATION: n = g, a, c o r t

US-09-223-040-1

Query Match 100.0%; Score 396; DB 4; Length 2287;

Best Local Similarity 100.0%; Pred. No. 2e-91;

Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACGCCGCGTCCGATAACTTCCAGCTGTCAGGTTGGCAGGATTCGCCATTCCCGATC 60

63 ACGCCGCGTCCGATAACTTCCAGCTGTCAGGTTGGCAGGATTCGCCATTCCCGATC 122

Qy 61 GGGCAGCGATGCGGATCGCGGCGAGATCGATCGGTTGGGGGTCCACCACCGTTTCAT 120

123 GGGCAGCGATGCGGATCGCGGCGAGATCGATCGGTTGGGGGTCCACCACCGTTTCAT 182

Qy 121 ATCCGGGCTACCGCCTTCCCTCGGCTTGGGTTGTCGACACACACGCGGCGCAGCA 180

Db	183	ATCGGGCGCTACGCCCTTCTCGGCTTGGGTGTGTGTCGACAAACAGCGCAACGGCGCACGA	242
Qy	181	GTCCAAACGGGTGGTCCGGAGCGCTCCGGCGGCAAGCTCTCGGCATCTCCACGGCGACGTG	240
Db	243	GTCCAAACGGGTGGTCCGGAGCGCTCCGGCGGCAAGCTCTCGGCATCTCCACGGCGACGTG	302
Qy	241	ATCACCGCGGTGCAGCGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGGCTTTAAC	300
Db	303	ATCACCGCGGTGCAGCGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGGCTTTAAC	362
Qy	301	GGGCATCATCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTCGGGCGGACCGCT	360
Db	363	GGGCATCATCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTCGGGCGGACCGCT	422
Qy	361	ACAGGGAACGTGACATTTGGCCGAGGAGCCCGCGGCC	396
Db	423	ACAGGGAACGTGACATTTGGCCGAGGAGCCCGCGGCC	458
RESULT 2			
US-09-103-840A-2			
; Sequence 2, Application US/09103840A			
; Patent No. 6294328			
; GENERAL INFORMATION:			
; APPLICANT: FLEISCHMAN, Robert D.			
; APPLICANT: WHITE, Owen R.			
; APPLICANT: FRASER, Claire M.			
; APPLICANT: VENTER, John C.			
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM			
; TITLE OF INVENTION: TUBERCULOSIS			
; FILE REFERENCE: 24366-20007.00			
; CURRENT APPLICATION NUMBER: US/09/103.840A			
; CURRENT FILING DATE: 1998-06-24			
; NUMBER OF SEQ ID NOS: 2			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 2			
; LENGTH: 4403765			
; TYPE: DNA			
; ORGANISM: Mycobacterium tuberculosis			
; FEATURE:			
; OTHER INFORMATION: CDC 1551			
; OTHER INFORMATION: "n" bases at various positions throughout the sequence			
; OTHER INFORMATION: represent a, t, c or g			
US-09-103-840A-2			
Query Match 100.0%; Score 396; DB 3; Length 4403765;			
Best Local Similarity 100.0%; Pred. No. 7e-91;			
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	ACGGCGCGGTCCGATTAACCTTCACAGTGTCCACAGGTTGGGCGAGGATTCGCCATTCGGATC	60
Db	151984	ACGGCGCGGTCCGATTAACCTTCACAGTGTCCACAGGTTGGGCGAGGATTCGCCATTCGGATC	152043
Qy	61	GGCGAGGGGATGGCGATCCGGGGCCAGATCCGATCGGTGGGGGTCAACCCACCGTTTCAT	120
Db	152044	GGCGAGGGGATGGCGATCCGGGGCCAGATCCGATCGGTGGGGGTCAACCCACCGTTTCAT	152103
Qy	121	ATCGGGCGCTACCGCTTCCTCGGTTGGGTGTTGTCGACAAACAGCGCAACGGCGCACGA	180
Db	152104	ATCGGGCGCTACCGCTTCCTCGGTTGGGTGTTGTCGACAAACAGCGCAACGGCGCACGA	152163
Qy	181	GTCCAAACGGGTGGTCCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGACGTG	240
Db	152164	GTCCAAACGGGTGGTCCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGACGTG	152223
Qy	241	ATCACCGCGGTGCAGCGGCGCTCCGATCAACTCGGCCACCGCGATGGCGACGGCTTTAAC	300
Db	152224	ATCACCGCGGTGCAGCGGCGCTCCGATCAACTCGGCCACCGCGATGGCGACGGCTTTAAC	152283
Qy	301	GGGCATCATCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTCGGGCGGACCGCT	360
Db	152284	GGGCATCATCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTCGGGCGGACCGCT	152343

;; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY  
;; NUMBER OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS  
;; NUMBER OF SEQUENCES: 153  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: SEED and BERRY LLP  
;; STREET: 6300 Columbia Center, 701 Fifth Avenue  
;; CITY: Seattle  
;; STATE: Washington  
;; COUNTRY: USA  
;; ZIP: 98104-7092

;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30

;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/818,112  
;; FILING DATE: 13-MAR-1997

;; CLASSIFICATION: 424  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Maki, David J.  
;; REGISTRATION NUMBER: 31,392

;; REFERENCE/DOCKET NUMBER: 210121.411C6  
;; TELEPHONE: (206) 622-4900  
;; TELEFAX: (206) 682-6031

;; INFORMATION FOR SEQ ID NO: 4:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 447 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear

;; US-08-818-112-4

Query Match 99.6%; Score 394.4; DB 3; Length 447;  
Best Local Similarity 99.7%; Pred. No. 3.8e-91;

Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACGGCGCGTCCGATTAATCCAGCTGCTCCAGGGTGGGAGGATTCCGATTCGATC 60  
Db 11 ACGGCGCGTCCGATTAATCCAGCTGCTCCAGGGTGGGAGGATTCCGATTCGATC 70

QY 61 GGGCAGCGATGGCGATCGGGCCAGATCCGATCGGGTGGGAGGATTCCGATTCGATC 120  
Db 71 GGGCAGCGATGGCGATCGGGCCAGATCCGATCGGGTGGGAGGATTCCGATTCGATC 130

QY 121 ATCGGCGCTACCGCTTCTCCGCTTGGGTGTTGTCACAAACGCGACGCGACGCA 180  
Db 131 ATCGGCGCTACCGCTTCTCCGCTTGGGTGTTGTCACAAACGCGACGCGACGCA 190

QY 181 GTCCAAACGCGTGTGCGGGAGCGCTCCGGGCGCAAGTCTCGCATCTCCACCGCGAGCTG 240  
Db 191 GTCCAAACGCGTGTGCGGGAGCGCTCCGGGCGCAAGTCTCGCATCTCCACCGCGAGCTG 250

QY 241 ATCCACCGCGTGCAGCGCGCTCCGATCAATCGGCCACCGCGATGGCGGCGCTTAAC 300  
Db 251 ATCCACCGCGTGCAGCGCGCTCCGATCAATCGGCCACCGCGATGGCGGCGCTTAAC 310

QY 301 GGGCATCATCCCGGTACGCTCATCTCGGTGACCTGGCAACCAAGTCGGGCGCGACGCT 360  
Db 311 GGGCATCATCCCGGTACGCTCATCTCGGTGACCTGGCAACCAAGTCGGGCGCGACGCT 370

QY 361 ACAGGAAACGTGACATTGGCCGAGGACCCCGGCC 396  
Db 371 ACAGGAAACGTGACATTGGCCGAGGACCCCGGCC 406

RESULT 5  
US-08-818-111-4  
; Sequence 4, Application US/08818111  
; Patent No. 6338852  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.

;; APPLICANT: Skeiky, Yasir A.W.  
;; APPLICANT: Dillon, Davin C.  
;; APPLICANT: Campos-Neto, Antonia  
;; APPLICANT: Houghton, Raymond  
;; APPLICANT: Vedvick, Thomas S.  
;; APPLICANT: Twardzik, Daniel R.  
;; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF  
;; NUMBER OF SEQUENCES: 148  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: SEED and BERRY LLP  
;; STREET: 6300 Columbia Center, 701 Fifth Avenue  
;; CITY: Seattle  
;; STATE: Washington  
;; COUNTRY: USA  
;; ZIP: 98104-7092

;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30

;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/818,111  
;; FILING DATE: 13-MAR-1997

;; CLASSIFICATION: 424  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Maki, David J.  
;; REGISTRATION NUMBER: 31,392

;; REFERENCE/DOCKET NUMBER: 210121.417C6  
;; TELEPHONE: (206) 622-4900  
;; TELEFAX: (206) 682-6031

;; INFORMATION FOR SEQ ID NO: 4:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 447 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear

;; US-08-818-111-4

Query Match 99.6%; Score 394.4; DB 4; Length 447;  
Best Local Similarity 99.7%; Pred. No. 3.8e-91;

Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACGGCGCGTCCGATTAATCCAGCTGCTCCAGGGTGGGAGGATTCCGATTCGATC 60  
Db 11 ACGGCGCGTCCGATTAATCCAGCTGCTCCAGGGTGGGAGGATTCCGATTCGATC 70

QY 61 GGGCAGCGATGGCGATCGGGCCAGATCCGATCGGGTGGGAGGATTCCGATTCGATC 120  
Db 71 GGGCAGCGATGGCGATCGGGCCAGATCCGATCGGGTGGGAGGATTCCGATTCGATC 130

QY 121 ATCGGCGCTACCGCTTCTCCGCTTGGGTGTTGTCACAAACGCGACGCGACGCA 180  
Db 131 ATCGGCGCTACCGCTTCTCCGCTTGGGTGTTGTCACAAACGCGACGCGACGCA 190

QY 181 GTCCAAACGCGTGTGCGGGAGCGCTCCGGGCGCAAGTCTCGCATCTCCACCGCGAGCTG 240  
Db 191 GTCCAAACGCGTGTGCGGGAGCGCTCCGGGCGCAAGTCTCGCATCTCCACCGCGAGCTG 250

QY 241 ATCCACCGCGTGCAGCGCGCTCCGATCAATCGGCCACCGCGATGGCGGCGCTTAAC 300  
Db 251 ATCCACCGCGTGCAGCGCGCTCCGATCAATCGGCCACCGCGATGGCGGCGCTTAAC 310

QY 301 GGGCATCATCCCGGTACGCTCATCTCGGTGACCTGGCAACCAAGTCGGGCGCGACGCT 360  
Db 311 GGGCATCATCCCGGTACGCTCATCTCGGTGACCTGGCAACCAAGTCGGGCGCGACGCT 370

QY 361 ACAGGAAACGTGACATTGGCCGAGGACCCCGGCC 396  
Db 371 ACAGGAAACGTGACATTGGCCGAGGACCCCGGCC 406

RESULT 6  
US-08-818-111-4  
; Sequence 4, Application US/08818111  
; Patent No. 6338852  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.





Db 311 GGGCATCATCCGGTGCAGTCATCTCGGTGAACATGGCAACAAAGTCGGGGCGGACGCGT 370  
Qy 361 ACAGGGAACGTGACATTGGCGAGGAGACCCCGGCC 396  
Db 371 ACAGGGAACGTGACATTGGCGAGGAGACCCCGGCC 406

RESULT 8  
US-08-818-112-17  
; Sequence 17, Application US/08818112  
; Patent No. 6290969  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Campos-Neto, Antonio  
; APPLICANT: Houghton, Raymond  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Twardzik, Daniel R.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS  
; NUMBER OF SEQUENCES: 153  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED and BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/818,112  
; FILING DATE: 13-MAR-1997  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maki, David J.  
; REGISTRATION NUMBER: 31,392  
; REFERENCE/DOCKET NUMBER: 210121.411C6  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1872 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-818-112-17

Query Match 98.8%; Score 391.2; DB 3; Length 1872;  
Best Local Similarity 99.2%; Pred. No. 3.1e-90;  
Matches 393; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ACGGCCGCGTCCGATAACTCCAGCTGTCCAGGGTGGGAGGATTCGCCATTCCGATC 60  
Db 758 ACGGCCGCGTCCGATAACTCCAGCTGTCCAGGGTGGGAGGATTCGCCATTCCGATC 817  
Qy 61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGTCAACCACCGTTTCAT 120  
Db 818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGTCAACCACCGTTTCAT 877  
Qy 121 ATCGGGCTACCGCTTCTCCCTCGGCTTGGGTGTTTCGACACACACGCGGACGCA 180  
Db 878 ATCGGGCTACCGCTTCTCCCTCGGCTTGGGTGTTTCGACACACACGCGGACGCA 937  
Qy 181 GTCCAAACGCGTGGTGGGAGGCGTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGAGTG 240  
Db 938 GTCCAAACGCGTGGTGGGAGGCGTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGAGTG 997

Qy 241 ATCACCGCGTGCAGCGCTCCGATCAATCGGCCACCGCGATGGGGAGCGCGCTTAAC 300  
Db 998 ATCACCGCGTGCAGCGCTCCGATCAATCGGCCACCGCGATGGGGAGCGCGCTTAAC 1057  
Qy 301 GGGCATCATCCCGTGCAGTCATCTCGGTGACCTGGCAAAACCAAGTCGGGGCGGACGCGT 360  
Db 1058 GGGCATCATCCCGTGCAGTCATCTCGGTGAACTGGCAAAACCAAGTCGGGGCGGACGCGT 1117  
Qy 361 ACAGGGAACGTGACATTGGCGAGGAGACCCCGGCC 396  
Db 1118 ACAGGGAACGTGACATTGGCGAGGAGACCCCGGCC 1153

RESULT 9  
US-08-818-111-17  
; Sequence 17, Application US/08818111  
; Patent No. 6338852  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Campos-Neto, Antonio  
; APPLICANT: Houghton, Raymond  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Twardzik, Daniel R.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF  
; NUMBER OF SEQUENCES: 148  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED and BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/818,111  
; FILING DATE: 13-MAR-1997  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maki, David J.  
; REGISTRATION NUMBER: 31,392  
; REFERENCE/DOCKET NUMBER: 210121.417C6  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1872 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-818-111-17

Query Match 98.8%; Score 391.2; DB 4; Length 1872;  
Best Local Similarity 99.2%; Pred. No. 3.1e-90;  
Matches 393; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ACGGCCGCGTCCGATAACTCCAGCTGTCCAGGGTGGGAGGATTCGCCATTCCGATC 60  
Db 758 ACGGCCGCGTCCGATAACTCCAGCTGTCCAGGGTGGGAGGATTCGCCATTCCGATC 817  
Qy 61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGTCAACCACCGTTTCAT 120  
Db 818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGTCAACCACCGTTTCAT 877  
Qy 121 ATCGGGCTACCGCTTCTCCCTCGGCTTGGGTGTTTCGACACACACGCGGACGCA 180  
Db 878 ATCGGGCTACCGCTTCTCCCTCGGCTTGGGTGTTTCGACACACACGCGGACGCA 937

QY 181 GTCCAAACGGGTGGTCCGGAGACGCTCCGGCGGCAAGTCTCGGCAATCTCCACCGGCGACGTG 240  
|||||  
Db 938 GTCCAAACGGGTGGTCCGGAGCGCTCCGGCGGCAAGTCTCGGCAATCTCCACCGGCGACGTG 997  
|||||  
QY 241 ATCACCAGGCGTCCGGCGGCTCCGATCACTCCGCGCCACCGCATGGCGGACGCGCTTAAC 300  
|||||  
Db 998 ATCACCAGGCGTCCGGCGGCTCCGATCACTCCGCGCCACCGCATGGCGGACGCGCTTAAC 1057  
|||||  
QY 301 GGGCATCATCCCGGTGACGTCTCATCTCGGTGACCTGGCAAAACCAAGTCGGGGCGGCGCGCT 360  
|||||  
Db 1058 GGGCATCATCCCGGTGACGTCTCATCTCGGTGACCTGGCAAAACCAAGTCGGGGCGGCGCGCT 1117  
|||||  
QY 361 ACAGGGAACGTGACATTTGGCCGAGGACCCCGGGCC 396  
|||||  
Db 1118 ACAGGGAACGTGACATTTGGCCGAGGACCCCGGGCC 1153  
|||||

## RESULT 10

US-09-056-556-17  
; Sequence 17, Application US/09056556  
; Patent No. 6350456  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Dillon, Davin C.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND  
; NUMBER OF SEQUENCES: 241  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED and BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/056,556  
; FILING DATE: 07-APR-1998  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maki, David J.  
; REGISTRATION NUMBER: 31,392  
; REFERENCE/DOCKET NUMBER: 210121.457  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1872 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-09-056-556-17

Query Match 98.8%; Score 391.2; DB 4; Length 1872;  
Best Local Similarity 99.2%; Pred. No. 3.1e-90;  
Matches 593; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 ACAGGCGGCGTCCGATAACTTCCAGCTGTCCAGGCGTCCAGGCGGCGGAGTTCGCCATTCGGATC 60  
|||||  
Db 758 ACAGGCGGCGTCCGATAACTTCCAGCTGTCCAGGCGTCCAGGCGGCGGAGTTCGCCATTCGGATC 817  
|||||  
QY 61 GGGCAGGCGATGGCGATCGCGGCGCAGATCCGATCGGTGGGGGTCCACCCACCGTTTAT 120  
|||||  
Db 818 GGGCAGGCGATGGCGATCGCGGCGCAGATCCGATCGGTGGGGGTCCACCCACCGTTTAT 877  
|||||  
QY 121 ATCGGGCGCTACGCGCTTCTCGGCTTGGTGTTCGACAAACCAACGCGGCGGCGACGA 180  
|||||

Db 878 ATCGGGCGCTACGCGCTTCTCGGCTTGGTGTTCGACAAACCAACGCGGCGGCGACGA 937  
|||||  
QY 181 GTCCAAACGGGTGGTCCGGAGCGCTCCGGCGGCAAGTCTCGGCAATCTCCACCGGCGACGTG 240  
|||||  
Db 938 GTCCAAACGGGTGGTCCGGAGCGCTCCGGCGGCAAGTCTCGGCAATCTCCACCGGCGACGTG 997  
|||||  
QY 241 ATCACCAGGCGTCCGGCGGCTCCGATCACTCCGCGCCACCGCATGGCGGACGCGCTTAAC 300  
|||||  
Db 998 ATCACCAGGCGTCCGGCGGCTCCGATCACTCCGCGCCACCGCATGGCGGACGCGCTTAAC 1057  
|||||  
QY 301 GGGCATCATCCCGGTGACGTCTCATCTCGGTGACCTGGCAAAACCAAGTCGGGGCGGCGCGCT 360  
|||||  
Db 1058 GGGCATCATCCCGGTGACGTCTCATCTCGGTGACCTGGCAAAACCAAGTCGGGGCGGCGCGCT 1117  
|||||  
QY 361 ACAGGGAACGTGACATTTGGCCGAGGACCCCGGGCC 396  
|||||  
Db 1118 ACAGGGAACGTGACATTTGGCCGAGGACCCCGGGCC 1153  
|||||

## RESULT 11

US-09-072-596-17  
; Sequence 17, Application US/09072596  
; Patent No. 6458366  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Campos-Neto, Antonia  
; APPLICANT: Houghton, Raymond  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Twardzik, Daniel R.  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Hendrickson, Ronald C.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF  
; NUMBER OF SEQUENCES: 350  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED and BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/072,596  
; FILING DATE: 05-MAY-1998  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maki, David J.  
; REGISTRATION NUMBER: 31,392  
; REFERENCE/DOCKET NUMBER: 210121.417C9  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1872 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-09-072-596-17

Query Match 98.8%; Score 391.2; DB 4; Length 1872;  
Best Local Similarity 99.2%; Pred. No. 3.1e-90;  
Matches 593; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 ACAGGCGGCGTCCGATAACTTCCAGCTGTCCAGGCGTCCAGGCGGCGGAGTTCGCCATTCGGATC 60  
|||||  
Db 758 ACAGGCGGCGTCCGATAACTTCCAGCTGTCCAGGCGTCCAGGCGGCGGAGTTCGCCATTCGGATC 817  
|||||

Qy 61 GGCAGCGGATGGCGATCGCGGGCCAGATCCGATCGGTTGGGGGTCAACCCACCGTTTCAT 120  
Db 818 GGCAGCGGATGGCGATCGCGGGCCAAATCCGATCGGTTGGGGGTCAACCCACCGTTTCAT 877  
Qy 121 ATCCGGGCTACCGCTTCCCTCGGCTTGGTGTTCGACAAACAGCGGCGGACGGA 180  
Db 878 ATCCGGGCTACCGCTTCCCTCGGCTTGGTGTTCGACAAACAGCGGCGGACGGA 937  
Qy 181 GTCCAAACCGGTGTCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGAGTG 240  
Db 938 GTCCAAACCGGTGTCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGAGTG 997  
Qy 241 ATCCAGCGGTGTCAGCGGCTCCGATCAACTCGGCCACCGCGATGGCGGCGCTTAAAC 300  
Db 998 ATCCAGCGGTGTCAGCGGCTCCGATCAACTCGGCCACCGCGATGGCGGCGCTTAAAC 1057  
Qy 301 GGGCATCATCCCGTGAGCTCATCTCGGTGACCTGGCAACCAAGTGGGGGGGACCGGT 360  
Db 1058 GGGCATCATCCCGTGAGCTCATCTCGGTGACCTGGCAACCAAGTGGGGGGGACCGGT 1117  
Qy 361 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 396  
Db 1118 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 1153

## RESULT 12

US-09-736-457-1862  
; Sequence 1862, Application US/09736457  
; Patent No. 6509448  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Lodes, Michael A.  
; APPLICANT: Fanger, Gary  
; APPLICANT: Vedvick, Tom  
; APPLICANT: Carter, Darrick  
; APPLICANT: Retter, Marc  
; APPLICANT: Mannion, Jane  
; APPLICANT: Fan, Liqun  
; APPLICANT: Wang, Aijun  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.478C15  
; CURRENT APPLICATION NUMBER: US/09/736.457  
; CURRENT FILING DATE: 2000-12-13  
; NUMBER OF SEQ ID NOS: 1864  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1862  
; LENGTH: 822  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-736-457-1862

Query Match 89.4%; Score 354; DB 4; Length 822;  
Best Local Similarity 95.7%; Pred. No. 6.8e-81;  
Matches 379; Conservative 0; Mismatches 5; Indels 12; Gaps 1;  
Qy 1 ACGCCGCGTCCGATTAACCTCCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCGGATC 60  
Db 22 ACGCCGCGTCCGATTAACCTCCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCGGATC 81  
Qy 61 GGGCAGCGATGGCGATCGCGGGCCAGATCCGATCGGTTGGGGGTCAACCCACCGTTTCAT 120  
Db 82 GGGCAGCGATGGCGATCGCGGGCCAGATC-----AAGCTTCCACCGTTTCAT 129  
Qy 121 ATCCGGGCTACCGCTTCCCTCGGCTTGGTGTTCGACAAACAGCGGCGGACGGA 180  
Db 130 ATCCGGGCTACCGCTTCCCTCGGCTTGGTGTTCGACAAACAGCGGCGGACGGA 189  
Qy 181 GTCCAAACCGGTGTCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGAGTG 240  
Db 190 GTCCAAACCGGTGTCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGAGTG 249

Qy 241 ATCCAGCGGTGTCAGCGGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAAC 300  
Db 250 ATCCAGCGGTGTCAGCGGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAAC 309  
Qy 301 GGGCATCATCCCGTGAGCTCATCTCGGTGACCTGGCAACCAAGTGGGGGGGACCGGT 360  
Db 310 GGGCATCATCCCGTGAGCTCATCTCGGTGACCTGGCAACCAAGTGGGGGGGACCGGT 369  
Qy 361 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 396  
Db 370 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 405

## RESULT 13

US-09-643-597-353  
; Sequence 353, Application US/09643597  
; Patent No. 6426072  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hosken, Nancy  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: McNeill, Patricia D.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.455C11  
; CURRENT APPLICATION NUMBER: US/09/643.597  
; CURRENT FILING DATE: 2000-08-21  
; NUMBER OF SEQ ID NOS: 369  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 353  
; LENGTH: 900  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-643-597-353

Query Match 89.4%; Score 354; DB 4; Length 900;  
Best Local Similarity 95.7%; Pred. No. 6.9e-81;  
Matches 379; Conservative 0; Mismatches 5; Indels 12; Gaps 1;  
Qy 1 ACGCCGCGTCCGATTAACCTCCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCGGATC 60  
Db 22 ACGCCGCGTCCGATTAACCTCCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCGGATC 81  
Qy 61 GGGCAGCGATGGCGATCGCGGGCCAGATCCGATCGGTTGGGGGTCAACCCACCGTTTCAT 120  
Db 82 GGGCAGCGATGGCGATCGCGGGCCAGATC-----AAGCTTCCACCGTTTCAT 129  
Qy 121 ATCCGGGCTACCGCTTCCCTCGGCTTGGTGTTCGACAAACAGCGGCGGACGGA 180  
Db 130 ATCCGGGCTACCGCTTCCCTCGGCTTGGTGTTCGACAAACAGCGGCGGACGGA 189  
Qy 181 GTCCAAACCGGTGTCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGAGTG 240  
Db 190 GTCCAAACCGGTGTCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGAGTG 249  
Qy 241 ATCCAGCGGTGTCAGCGGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAAC 300  
Db 250 ATCCAGCGGTGTCAGCGGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAAC 309  
Qy 301 GGGCATCATCCCGTGAGCTCATCTCGGTGACCTGGCAACCAAGTGGGGGGGACCGGT 360  
Db 310 GGGCATCATCCCGTGAGCTCATCTCGGTGACCTGGCAACCAAGTGGGGGGGACCGGT 369  
Qy 361 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 396  
Db 1118 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 1153

Db 370 ACAGGGAACGTGACATTTGGCCGAGGACCCCGGCC 405

## RESULT 14

US-09-606-421B-353

; Sequence 353, Application US/09606421B

; Patent No. 6531315

; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong

; APPLICANT: Fan, Liqun

; APPLICANT: Kalos, Michael D.

; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: Hosken, Nancy

; APPLICANT: Fanger, Gary R.

; APPLICANT: Li, Samuel X.

; APPLICANT: Wang, Aljun

; APPLICANT: Skeiky, Yasir A.W.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER

; FILE REFERENCE: 210121.455C9

; CURRENT APPLICATION NUMBER: US/09/606,421B

; CURRENT FILING DATE: 2000-06-28

; NUMBER OF SEQ ID NOS: 358

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 353

; LENGTH: 900

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-606-421B-353

Query Match 89.4%; Score 354; DB 4; Length 900;

Best Local Similarity 95.7%; Pred. No. 6.9e-81;

Matches 379; Conservative 0; Mismatches 5; Indels 12; Gaps 1;

QY 1 ACGGCGCGCTCCGATAACTTCCAGCTGTCCAGGCTGGCGAGGATTCGCCATTCCGATC 60

Db 22 ACGGCGCGCTCCGATAACTTCCAGCTGTCCAGGCTGGCGAGGATTCGCCATTCCGATC 81

QY 61 GGGCAGGCGATGGCGATCGCGGCCAGATCCGATCGGTGGGGGTCAACCCACCGTTTCA 120

Db 82 GGGCAGGCGATGGCGATCGCGGCCAGATC-----AAGCTTCCCACCGTTTCA 129

QY 121 ATCGGGCTTACCGCTTCTCGGCTTGGGTGTGTCGACAAACGCGGACGCGACGCA 180

Db 130 ATCGGGCTTACCGCTTCTCGGCTTGGGTGTGTCGACAAACGCGGACGCGACGCA 189

QY 181 GTCCAAACGCGTGGTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGGACG 240

Db 190 GTCCAAACGCGTGGTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGGACG 249

QY 241 ATCACCGGGTTCGACGGGCTCCGATCAACTCGGCCACCGGATGGCGGACGCGCTTAAC 300

Db 250 ATCACCGGGTTCGACGGGCTCCGATCAACTCGGCCACCGGATGGCGGACGCGCTTAAC 309

QY 301 GGGCATATCCCGGTGACGTCTCGGTGACCTGGCAACCAAGTCGGGGGCGACGCGT 360

Db 310 GGGCATATCCCGGTGACGTCTCGGTGACCTGGCAACCAAGTCGGGGGCGACGCGT 369

QY 361 ACAGGGAACGTGACATTTGGCCGAGGACCCCGGCC 396

Db 370 ACAGGGAACGTGACATTTGGCCGAGGACCCCGGCC 405

## RESULT 15

US-09-736-457-1861

; Sequence 1861, Application US/09736457

; Patent No. 6509448

; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong

; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: Lodes, Michael A.

; APPLICANT: Fanger, Gary

; APPLICANT: Vedvick, Tom

; APPLICANT: Carter, Darrick

; APPLICANT: Retter, Marc

; APPLICANT: Mannion, Jane

; APPLICANT: Fan, Liqun

; APPLICANT: Wang, Aljun

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; FILE OF INVENTION: DIAGNOSIS OF LUNG CANCER

; FILE REFERENCE: 210121.478C15

; CURRENT APPLICATION NUMBER: US/09/736,457

; CURRENT FILING DATE: 2000-12-13

; NUMBER OF SEQ ID NOS: 1864

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 1861

; LENGTH: 945

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-736-457-1861

Query Match 89.4%; Score 354; DB 4; Length 945;

Best Local Similarity 95.7%; Pred. No. 7e-81;

Matches 379; Conservative 0; Mismatches 5; Indels 12; Gaps 1;

QY 1 ACGGCGCGCTCCGATAACTTCCAGCTGTCCAGGCTGGCGAGGATTCGCCATTCCGATC 60

Db 22 ACGGCGCGCTCCGATAACTTCCAGCTGTCCAGGCTGGCGAGGATTCGCCATTCCGATC 81

QY 61 GGGCAGGCGATGGCGATCGCGGCCAGATCCGATCGGTGGGGGTCAACCCACCGTTTCA 120

Db 82 GGGCAGGCGATGGCGATCGCGGCCAGATC-----AAGCTTCCCACCGTTTCA 129

QY 121 ATCGGGCTTACCGCTTCTCGGCTTGGGTGTGTCGACAAACGCGGACGCGGACGCA 180

Db 130 ATCGGGCTTACCGCTTCTCGGCTTGGGTGTGTCGACAAACGCGGACGCGGACGCA 189

QY 181 GTCCAAACGCGTGGTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGGACG 240

Db 190 GTCCAAACGCGTGGTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGGACG 249

QY 241 ATCACCGGGTTCGACGGGCTCCGATCAACTCGGCCACCGGATGGCGGACGCGCTTAAC 300

Db 250 ATCACCGGGTTCGACGGGCTCCGATCAACTCGGCCACCGGATGGCGGACGCGCTTAAC 309

QY 301 GGGCATATCCCGGTGACGTCTCGGTGACCTGGCAACCAAGTCGGGGGCGACGCGT 360

Db 310 GGGCATATCCCGGTGACGTCTCGGTGACCTGGCAACCAAGTCGGGGGCGACGCGT 369

QY 361 ACAGGGAACGTGACATTTGGCCGAGGACCCCGGCC 396

Db 370 ACAGGGAACGTGACATTTGGCCGAGGACCCCGGCC 405

Search completed: September 5, 2003, 08:28:53

Job time : 87 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2003, 07:20:04 ; Search time 1535 Seconds  
(without alignments)  
593.349 Million cell updates/sec

Title: US-09-684-215A-3  
Perfect score: 396  
Sequence: 1 acggcgcgtccgataactt.....tgccgagggaccgccgccc 396

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1537136 seqs, 1149988732 residues

Total number of hits satisfying chosen parameters: 3074272

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_NA:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
16: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
17: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	396	100.0	702	9	US-09-387-849-27
2	396	100.0	702	12	US-10-359-460-27
3	396	100.0	1068	10	US-09-712-363-15
4	396	100.0	2287	9	US-09-287-849-1
5	396	100.0	2287	12	US-10-359-460-1
6	394.4	99.6	447	12	US-10-084-843-4
7	394.4	99.6	447	12	US-10-193-002-4
8	391.2	98.8	1872	12	US-10-084-843-17
9	391.2	98.8	1872	12	US-10-193-002-17
10	354	89.4	675	9	US-09-759-143-822
11	354	89.4	675	9	US-09-780-669-822
12	354	89.4	675	10	US-09-822-827-822
13	354	89.4	675	10	US-09-895-793-822
14	354	89.4	675	10	US-09-895-814-822
15	354	89.4	675	12	US-10-144-678A-822
16	354	89.4	675	13	US-10-012-896-822

17	354	89.4	822	10	US-09-736-457-1862	Sequence 1862, Ap
18	354	89.4	822	10	US-09-902-941-1862	Sequence 1862, Ap
19	354	89.4	822	10	US-09-849-626-1862	Sequence 1862, Ap
20	354	89.4	822	14	US-10-017-754-1862	Sequence 1862, Ap
21	354	89.4	900	9	US-09-735-705-353	Sequence 353, App
22	354	89.4	900	10	US-09-850-716A-353	Sequence 353, App
23	354	89.4	900	10	US-09-897-778-353	Sequence 353, App
24	354	89.4	900	12	US-10-117-982-353	Sequence 353, App
25	354	89.4	915	9	US-09-759-143-834	Sequence 834, App
26	354	89.4	915	9	US-09-780-669-834	Sequence 834, App
27	354	89.4	915	9	US-09-822-827-834	Sequence 834, App
28	354	89.4	915	10	US-09-895-793-834	Sequence 834, App
29	354	89.4	915	10	US-09-895-814-834	Sequence 834, App
30	354	89.4	915	12	US-10-144-678A-834	Sequence 834, App
31	354	89.4	915	13	US-10-012-896-834	Sequence 834, App
32	354	89.4	945	10	US-09-736-457-1861	Sequence 1861, Ap
33	354	89.4	945	10	US-09-902-941-1861	Sequence 1861, Ap
34	354	89.4	945	10	US-09-849-626-1861	Sequence 1861, Ap
35	354	89.4	945	14	US-10-017-754-1861	Sequence 1861, Ap
36	354	89.4	1012	9	US-09-735-705-351	Sequence 351, App
37	354	89.4	1012	10	US-09-850-716A-351	Sequence 351, App
38	354	89.4	1012	10	US-09-897-778-351	Sequence 351, App
39	354	89.4	1012	12	US-10-117-982-351	Sequence 351, App
40	354	89.4	1035	9	US-09-922-217-1084	Sequence 1084, Ap
41	354	89.4	1035	10	US-09-833-263-1084	Sequence 1084, Ap
42	354	89.4	1035	11	US-09-938-864-388	Sequence 388, App
43	354	89.4	1035	13	US-10-025-380-1084	Sequence 1084, Ap
44	354	89.4	1035	14	US-10-125-635A-388	Sequence 388, App
45	354	89.4	1035	14	US-10-002-603-388	Sequence 388, App

ALIGNMENTS

RESULT 1

US-09-287-849-27  
; Sequence 27, Application US/09287849  
; Patent No. US20020009459A1  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Alderson, Mark  
; APPLICANT: Campos-Neto, Antonio  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Fusion proteins of Mycobacterium tuberculosis Antigens  
; TITLE OF INVENTION: and Their Uses  
; FILE REFERENCE: 014058-009020US  
; CURRENT APPLICATION NUMBER: US/09/287,849  
; CURRENT FILING DATE: 1999-04-07  
; PRIOR APPLICATION NUMBER: US 08/818,112  
; PRIOR FILING DATE: 1997-03-13  
; PRIOR APPLICATION NUMBER: US 08/942,578  
; PRIOR FILING DATE: 1997-10-01  
; PRIOR APPLICATION NUMBER: US 09/025,197  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 09/056,556  
; PRIOR FILING DATE: 1998-04-07  
; PRIOR APPLICATION NUMBER: US 09/223,040  
; PRIOR FILING DATE: 1998-12-30  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 27  
; LENGTH: 702  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:bi-fusion  
; OTHER INFORMATION: protein Ra12-DPPD (designated Mtb24), reading  
; OTHER INFORMATION: frame 1  
; NAME/KEY: CDS  
; LOCATION: (1)..(693)  
; OTHER INFORMATION: bi-fusion protein Ra12-DPPD (designated Mtb24),

OTHER INFORMATION: reading frame 1  
NAME/KEY: CDS  
LOCATION: (2)..(700)  
OTHER INFORMATION: reading frame 2  
NAME/KEY: CDS  
LOCATION: (3)..(701)  
OTHER INFORMATION: reading frame 3  
US-09-287-849-27

Query Match 100.0%; Score 396; DB 9; Length 702;  
Best Local Similarity 100.0%; Pred. No. 7.1e-104; Indels 0; Gaps 0;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGGTGGGCGAGGATTCCGCCATTCCGATC 60  
DB 25 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGGTGGGCGAGGATTCCGCCATTCCGATC 84  
QY 61 GGGCAGGCGATGGCGATCGGGCCAGATCCGATCGGGTGGGGTGCACCCACCGTTTAT 120  
DB 85 GGGCAGGCGATGGCGATCGGGCCAGATCCGATCGGGTGGGGTGCACCCACCGTTTAT 144  
QY 121 ATCGGGCTACCGCTTCTCGGCTTGGTGTGTCGACAAACGCAACGGCGCACGA 180  
DB 145 ATCGGGCTACCGCTTCTCGGCTTGGTGTGTCGACAAACGCAACGGCGCACGA 204  
QY 181 GTCCAAACGGTGGTCCGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGCGACGTG 240  
DB 205 GTCCAAACGGTGGTCCGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGCGACGTG 264  
QY 241 ATCCAGCGGTGCGAGCGCTCCGATCACTCGGCGCACCGGATCGGCGACGCTTAC 300  
DB 265 ATCCAGCGGTGCGAGCGCTCCGATCACTCGGCGCACCGGATCGGCGACGCTTAC 324  
QY 301 GGGCATCATCCCGGTGAGCTCATCTCGGTGAGCTGGCAACCAAGTCGGCGGCGACGGT 360  
DB 325 GGGCATCATCCCGGTGAGCTCATCTCGGTGAGCTGGCAACCAAGTCGGCGGCGACGGT 384  
QY 361 ACAGGAGCGTGACATTTGGCGAGGACCCCGGCC 396  
DB 385 ACAGGAGCGTGACATTTGGCGAGGACCCCGGCC 420

## RESULT 2

US-10-359-460-27  
Sequence 27, Application US/10359460  
Publication No. US2003014791A1  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Skelky, Yasir A.W.  
APPLICANT: Dillon, Davin C.  
APPLICANT: Alderson, Mark  
APPLICANT: Campos-Neto, Antonio  
APPLICANT: Corixa Corporation  
TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens  
FILE REFERENCE: 014058-009020US  
CURRENT APPLICATION NUMBER: US/10/359,460  
CURRENT FILING DATE: 2003-02-05  
PRIOR APPLICATION NUMBER: US/09/287,849  
PRIOR FILING DATE: 1999-04-07  
PRIOR APPLICATION NUMBER: US 08/818,112  
PRIOR FILING DATE: 1997-03-13  
PRIOR APPLICATION NUMBER: US 08/942,578  
PRIOR FILING DATE: 1997-10-01  
PRIOR APPLICATION NUMBER: US 09/025,197  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 09/056,556  
PRIOR FILING DATE: 1998-04-07  
PRIOR APPLICATION NUMBER: US 09/223,040  
PRIOR FILING DATE: 1998-12-30  
NUMBER OF SEQ ID NOS: 46  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 27

LENGTH: 702  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: bi-fusion  
OTHER INFORMATION: protein Ral2-DPPD (designated Mtb24), reading  
OTHER INFORMATION: frame 1  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(693)  
OTHER INFORMATION: bi-fusion protein Ral2-DPPD (designated Mtb24),  
OTHER INFORMATION: reading frame 1  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (2)..(700)  
OTHER INFORMATION: reading frame 2  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (3)..(701)  
OTHER INFORMATION: reading frame 3  
US-10-359-460-27

Query Match 100.0%; Score 396; DB 12; Length 702;  
Best Local Similarity 100.0%; Pred. No. 7.1e-104; Indels 0; Gaps 0;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGGTGGGCGAGGATTCCGCCATTCCGATC 60  
DB 25 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGGTGGGCGAGGATTCCGCCATTCCGATC 84  
QY 61 GGGCAGGCGATGGCGATCGGGCCAGATCCGATCGGGTGGGGTGCACCCACCGTTTAT 120  
DB 85 GGGCAGGCGATGGCGATCGGGCCAGATCCGATCGGGTGGGGTGCACCCACCGTTTAT 144  
QY 121 ATCGGGCTACCGCTTCTCGGCTTGGTGTGTCGACAAACGCAACGGCGCACGA 180  
DB 145 ATCGGGCTACCGCTTCTCGGCTTGGTGTGTCGACAAACGCAACGGCGCACGA 204  
QY 181 GTCCAAACGGTGGTCCGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGCGACGTG 240  
DB 205 GTCCAAACGGTGGTCCGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGCGACGTG 264  
QY 241 ATCCAGCGGTGCGAGCGCTCCGATCACTCGGCGCACCGGATCGGCGACGCTTAC 300  
DB 265 ATCCAGCGGTGCGAGCGCTCCGATCACTCGGCGCACCGGATCGGCGACGCTTAC 324  
QY 301 GGGCATCATCCCGGTGAGCTCATCTCGGTGAGCTGGCAACCAAGTCGGCGGCGACGGT 360  
DB 325 GGGCATCATCCCGGTGAGCTCATCTCGGTGAGCTGGCAACCAAGTCGGCGGCGACGGT 384  
QY 361 ACAGGAGCGTGACATTTGGCGAGGACCCCGGCC 396  
DB 385 ACAGGAGCGTGACATTTGGCGAGGACCCCGGCC 420

## RESULT 3

US-09-712-363-15  
Sequence 15, Application US/09712363  
Patent No. US20020164588A1  
GENERAL INFORMATION:  
APPLICANT: Elsenberg, David  
APPLICANT: Rotstein, Sergio H.  
APPLICANT: Marcotte, Edward M.  
TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND  
INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS  
FILE REFERENCE: 07419-032001  
CURRENT APPLICATION NUMBER: US/09/712,363  
CURRENT FILING DATE: 2000-11-13  
PRIOR APPLICATION NUMBER: PCT/US00/02246  
PRIOR FILING DATE: 2000-01-28  
PRIOR APPLICATION NUMBER: 60/179,531  
PRIOR FILING DATE: 2000-02-01  
PRIOR APPLICATION NUMBER: 60/117,844

;; PRIOR FILING DATE: 1999-01-29  
;; PRIOR APPLICATION NUMBER: 60/118,206,  
;; PRIOR FILING DATE: 1999-02-01  
;; PRIOR APPLICATION NUMBER: 60/126,593  
;; PRIOR FILING DATE: 1999-03-26  
;; PRIOR APPLICATION NUMBER: 60/134,093  
;; PRIOR FILING DATE: 1999-05-14  
;; PRIOR APPLICATION NUMBER: 60/134,092  
;; PRIOR FILING DATE: 1999-05-14  
;; PRIOR APPLICATION NUMBER: 60/165,124  
;; PRIOR FILING DATE: 1999-11-12  
;; PRIOR APPLICATION NUMBER: 60/165,086  
;; PRIOR FILING DATE: 1999-11-12  
;; NUMBER OF SEQ ID NOS: 292  
;; SEQ ID NO 15  
;; LENGTH: 1068  
;; TYPE: DNA  
;; ORGANISM: Mycobacterium tuberculosis  
US-09-712-363-15

Query Match 100.0%; Score 396; DB 10; Length 1068;  
Best Local Similarity 100.0%; Pred. No. 7.1e-104;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ACGCCCGCTCCGATAACTTCCAGCTGTCCAGGGTGGCGAGGATTCCGCAATTCGCGATC 60  
Db 670 ACGCCCGCTCCGATAACTTCCAGCTGTCCAGGGTGGCGAGGATTCCGCAATTCGCGATC 729  
Qy 61 GGGCAGCGATGGCGATCGGGCCAGATCGGATCGGGTGGGGGTCAACCCACCGTTTCAT 120  
Db 730 GGGCAGCGATGGCGATCGGGCCAGATCGGATCGGGTGGGGGTCAACCCACCGTTTCAT 789  
Qy 121 ATCGGGCTACCGCTTCTCCGCTTGGTGTCTCGACACAAACGCGCAACGCGGCACGA 180  
Db 790 ATCGGGCTACCGCTTCTCCGCTTGGTGTCTCGACACAAACGCGCAACGCGGCACGA 849  
Qy 181 GTCCAAACCGGTGTCGGAGCGCTCCGGCGCAAGTCTCGGCAATCTCCACCGGGACGTG 240  
Db 850 GTCCAAACCGGTGTCGGAGCGCTCCGGCGCAAGTCTCGGCAATCTCCACCGGGACGTG 909  
Qy 241 ATCACCAGCGTTCGAGCGGCTCCGATCAACTCGGCCACCGGATGGGGCGGCGCTTAAC 300  
Db 910 ATCACCAGCGTTCGAGCGGCTCCGATCAACTCGGCCACCGGATGGGGCGGCGCTTAAC 969  
Qy 301 GGGCATATCCCGTGACGTCTCGTGACCTGGCAACCAAGTCCGGCGGCGCACGCT 360  
Db 970 GGGCATATCCCGTGACGTCTCGTGACCTGGCAACCAAGTCCGGCGGCGCACGCT 1029  
Qy 361 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 396  
Db 1030 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 1065

RESULT 4  
US-09-287-849-1  
; Sequence 1, Application US/09287849  
; Patent No. US20020094591  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Alderson, Mark  
; APPLICANT: Campos-Neto, Antonio  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens  
; FILE REFERENCE: 014058-009020US  
; CURRENT FILING DATE: 1999-04-07  
; PRIOR FILING DATE: 1999-04-07  
; PRIOR APPLICATION NUMBER: US 08/818,112  
; PRIOR FILING DATE: 1997-03-13  
; PRIOR APPLICATION NUMBER: US 08/942,578

;; PRIOR FILING DATE: 1997-10-01  
;; PRIOR APPLICATION NUMBER: US 09/025,197  
;; PRIOR FILING DATE: 1998-02-18  
;; PRIOR APPLICATION NUMBER: US 09/056,556  
;; PRIOR FILING DATE: 1998-04-07  
;; PRIOR APPLICATION NUMBER: US 09/223,040  
;; PRIOR FILING DATE: 1998-12-30  
;; NUMBER OF SEQ ID NOS: 46  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 1  
;; LENGTH: 2287  
;; TYPE: DNA  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: Description of Artificial Sequence: tri-fusion  
;; OTHER INFORMATION: protein Rai2-TbH9-Ra35 (designated Mtb32-Mtb39)  
;; OTHER INFORMATION: fusion)  
;; NAME/KEY: modified\_base  
;; LOCATION: (30)  
;; OTHER INFORMATION: n = g, a, c or t  
;; NAME/KEY: modified\_base  
;; LOCATION: (33)  
;; OTHER INFORMATION: n = g, a, c or t  
;; NAME/KEY: CDS  
;; LOCATION: (42)..(2231)  
;; NAME/KEY: modified\_base  
;; LOCATION: (2270)  
;; OTHER INFORMATION: n = g, a, c or t  
US-09-287-849-1

Query Match 100.0%; Score 396; DB 9; Length 2287;  
Best Local Similarity 100.0%; Pred. No. 7.1e-104;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ACGCCCGCTCCGATAACTTCCAGCTGTCCAGGGTGGCGAGGATTCCGCAATTCGCGATC 60  
Db 63 ACGCCCGCTCCGATAACTTCCAGCTGTCCAGGGTGGCGAGGATTCCGCAATTCGCGATC 122  
Qy 61 GGGCAGCGATGGCGATCGGGCCAGATCGGATCGGGTGGGGGTCAACCCACCGTTTCAT 120  
Db 123 GGGCAGCGATGGCGATCGGGCCAGATCGGATCGGGTGGGGGTCAACCCACCGTTTCAT 182  
Qy 121 ATCGGGCTACCGCTTCTCCGCTTGGTGTCTCGACACAAACGCGCAACGCGGCACGA 180  
Db 183 ATCGGGCTACCGCTTCTCCGCTTGGTGTCTCGACACAAACGCGCAACGCGGCACGA 242  
Qy 181 GTCCAAACCGGTGTCGGAGCGCTCCGGCGCAAGTCTCGCATCTCCACCGGCGACGTG 240  
Db 243 GTCCAAACCGGTGTCGGAGCGCTCCGGCGCAAGTCTCGCATCTCCACCGGCGACGTG 302  
Qy 241 ATCACCAGCGTTCGAGCGGCTCCGATCAACTCGGCCACCGGATGGGGCGGCGCTTAAC 300  
Db 303 ATCACCAGCGTTCGAGCGGCTCCGATCAACTCGGCCACCGGATGGGGCGGCGCTTAAC 362  
Qy 301 GGGCATATCCCGTGACGTCTCGTGACCTGGCAACCAAGTCCGGCGGCGCACGCT 360  
Db 363 GGGCATATCCCGTGACGTCTCGTGACCTGGCAACCAAGTCCGGCGGCGCACGCT 422  
Qy 361 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 396  
Db 423 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 458

RESULT 5  
US-10-359-460-1  
; Sequence 1, Application US/10359460  
; Publication No. US20030147911A1  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Alderson, Mark  
; APPLICANT: Campos-Neto, Antonio

APPLICANT: Corixa Corporation  
TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens  
and Their Uses  
FILE REFERENCE: 014058-009020US  
CURRENT APPLICATION NUMBER: US/10/359,460  
CURRENT FILING DATE: 2003-02-05  
PRIOR APPLICATION NUMBER: US/09/287,849  
PRIOR FILING DATE: 1999-04-07  
PRIOR APPLICATION NUMBER: US 08/818,112  
PRIOR FILING DATE: 1997-03-13  
PRIOR APPLICATION NUMBER: US 08/942,578  
PRIOR FILING DATE: 1997-10-01  
PRIOR APPLICATION NUMBER: US 09/025,197  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 09/056,556  
PRIOR FILING DATE: 1998-04-07  
PRIOR APPLICATION NUMBER: US 09/223,040  
PRIOR FILING DATE: 1998-12-30  
NUMBER OF SEQ ID NOS: 46  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 2287  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: tri-fusion  
OTHER INFORMATION: protein Ra12-TbH9-Ra35 (designated Mtb32-Mtb39  
OTHER INFORMATION: fusion)  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (30)  
OTHER INFORMATION: n = g, a, c o r t  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (33)  
OTHER INFORMATION: n = g, a, c o r t  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (42)..(2231)  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (2270)  
OTHER INFORMATION: n = g, a, c o r t  
US-10-359-460-1

Query Match 100.0%; Score 396; DB 12; Length 2287;  
Best Local Similarity 100.0%; Pred. No. 7.1e-104;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGGGCGGTCGATAAATTCAGCTGCCAGGTGGCGAGGATTCGCCATTCGGATC 60  
DB 63 ACGGGCGGTCGATAAATTCAGCTGCCAGGTGGCGAGGATTCGCCATTCGGATC 122  
QY 61 GGGCAGGCGATGCCATCGCGGCCAGATCCAGTGGGTGGGGTCAACCCACCTTAT 120  
DB 123 GGGCAGGCGATGCCATCGCGGCCAGATCCAGTGGGTGGGGTCAACCCACCTTAT 182  
QY 121 ATCGGGCTACCGCTTCTCGGCTTGGGTGTGTCGACAAACGCAACGGCGCAGCA 180  
DB 183 ATCGGGCTACCGCTTCTCGGCTTGGGTGTGTCGACAAACGCAACGGCGCAGCA 242  
QY 181 GTCCAAACGGGTGGTGGGAGCGCTCCGGCGGCAAGTCTCGGATCTCCACCGGCGACGTG 240  
DB 243 GTCCAAACGGGTGGTGGGAGCGCTCCGGCGGCAAGTCTCGGATCTCCACCGGCGACGTG 302  
QY 241 ATCACCAGGCGTCCAGCGGCTCCGATCACTCGGCCACCGGATCGCGGCGCTTAAC 300  
DB 303 ATCACCAGGCGTCCAGCGGCTCCGATCACTCGGCCACCGGATCGCGGCGCTTAAC 362  
QY 301 GGGCATCATCCCGGTGAGCTCATCTCGGTGACCTGGCAAAACCAAGTCCGGCGGCGACGGT 360  
DB 363 GGGCATCATCCCGGTGAGCTCATCTCGGTGACCTGGCAAAACCAAGTCCGGCGGCGACGGT 422

QY 361 ACAGGGAACGACATGCGCGAGGAGGACCCCGGCC 396  
DB 423 ACAGGGAACGACATGCGCGAGGAGGACCCCGGCC 458

RESULT 6  
US-10-084-843-4  
Sequence 4, Application US/10084843  
Publication No. US20030143243A1  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
Skelley, Yasir A.W.  
Dillon, Davin C.  
Campos-Neto, Antonio  
Houghton, Raymond  
Vedvick, Thomas S.  
Twardzik, Daniel R.  
Lodes, Michael J.  
Hendrickson, Ronald C.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY  
AND DIAGNOSIS OF TUBERCULOSIS  
NUMBER OF SEQUENCES: 355  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED AND BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/084,843  
FILING DATE: 25-Feb-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/072,967  
FILING DATE: 05-MAY-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.411C9  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 447; base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-10-084-843-4

Query Match 99.6%; Score 394.4; DB 12; Length 447;  
Best Local Similarity 99.7%; Pred. No. 2e-103;  
Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACGGGCGGTCGATAAATTCAGCTGCCAGGTGGCGAGGATTCGCCATTCGGATC 60  
DB 11 ACGGGCGGTCGATAAATTCAGCTGCCAGGTGGCGAGGATTCGCCATTCGGATC 70  
QY 61 GGGCAGGCGATGCCATCGCGGCCAGATCCAGTGGGTGGGGTCAACCCACCTTAT 120  
DB 71 GGGCAGGCGATGCCATCGCGGCCAGATCCAGTGGGTGGGGTCAACCCACCTTAT 130  
QY 121 ATCGGGCTACCGCTTCTCGGCTTGGGTGTGTCGACAAACGCAACGGCGCAGCA 180  
DB 131 ATCGGGCTACCGCTTCTCGGCTTGGGTGTGTCGACAAACGCAACGGCGCAGCA 190





```
;
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 1872 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-10-084-843-17

Query Match          98.8%; Score 391.2; DB 12; Length 1872;
Best Local Similarity 99.2%; Pred. No. 1.7e-102;
Matches 393; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1  ACGGCGGCTCCGATAACTTCCAGCTGTCCAGGTGGCGAGGGATTCCGCCATTCGGATC 60
DB 758  ACGGCGGCTCCGATAACTTCCAGCTGTCCAGGTGGCGAGGGATTCCGCCATTCGGATC 817
QY 61  GGGCAGGCGATGGCATCGCGGCGCAGATCCGATCGGTGGGGTGCACCCACCGTTTCA 120
DB 818  GGGCAGGCGATGGCATCGCGGCGCAGATCCGATCGGTGGGGTGCACCCACCGTTTCA 877
QY 121  ATCGGCGCTACCGCTTCTCTCGGCTTGGGTGTGTCGACAAACGCGGCGCAGCA 180
DB 878  ATCGGCGCTACCGCTTCTCTCGGCTTGGGTGTGTCGACAAACGCGGCGCAGCA 937
QY 181  GTCCAAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
DB 938  GTCCAAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997
QY 241  ATCACCGCGTGCAGCGGCTCCGATCAACTCGGCGCAGCGCATCGGCGGCGCTTAA 300
DB 998  ATCACCGCGTGCAGCGGCTCCGATCAACTCGGCGCAGCGCATCGGCGGCGCTTAA 1057
QY 301  GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTTCGGGCGGCGACGGT 360
DB 1058  GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTTCGGGCGGCGACGGT 1117
QY 361  ACAGGGAAGTGACATTGGCCGAGGACCCCGGCC 396
DB 1118  ACAGGGAAGTGACATTGGCCGAGGACCCCGGCC 1153

RESULT 9
US-10-193-002-17
; Sequence 17, Application US/10193002
; Publication No. US20030135026A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
;           Skeiky, Yasir A.W.
;           Dillon, Davin C.
;           Campos-Neto, Antonia
;           Houghton, Raymond
;           Vedrick, Thomas S.
;           Twardzik, Daniel R.
;           Lodes, Michael J.
;           Hendrickson, Ronald C.
;
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
;                   TUBERCULOSIS
;
; NUMBER OF SEQUENCES: 350
; CORRESPONDENCE ADDRESS:
; ADDRESSER: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/193,002
; FILING DATE: 10-Jul-2002
```

```
;
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US/09/072,596
;   FILING DATE: 05-MAY-1998
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Maki, David J.
;     REGISTRATION NUMBER: 31,392
;     REFERENCE/DOCKET NUMBER: 210121.417C9
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (206) 622-4900
;     TELEFAX: (206) 682-6031
;
; INFORMATION FOR SEQ ID NO: 17:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 1872 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-10-193-002-17

Query Match          98.8%; Score 391.2; DB 12; Length 1872;
Best Local Similarity 99.2%; Pred. No. 1.7e-102;
Matches 393; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1  ACGGCGGCTCCGATAACTTCCAGCTGTCCAGGTGGCGAGGGATTCCGCCATTCGGATC 60
DB 758  ACGGCGGCTCCGATAACTTCCAGCTGTCCAGGTGGCGAGGGATTCCGCCATTCGGATC 817
QY 61  GGGCAGGCGATGGCATCGCGGCGCAGATCCGATCGGTGGGGTGCACCCACCGTTTCA 120
DB 818  GGGCAGGCGATGGCATCGCGGCGCAGATCCGATCGGTGGGGTGCACCCACCGTTTCA 877
QY 121  ATCGGCGCTACCGCTTCTCTCGGCTTGGGTGTGTCGACAAACGCGGCGCAGCA 180
DB 878  ATCGGCGCTACCGCTTCTCTCGGCTTGGGTGTGTCGACAAACGCGGCGCAGCA 937
QY 181  GTCCAAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
DB 938  GTCCAAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997
QY 241  ATCACCGCGTGCAGCGGCTCCGATCAACTCGGCGCAGCGCATCGGCGGCGCTTAA 300
DB 998  ATCACCGCGTGCAGCGGCTCCGATCAACTCGGCGCAGCGCATCGGCGGCGCTTAA 1057
QY 301  GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTTCGGGCGGCGACGGT 360
DB 1058  GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTTCGGGCGGCGACGGT 1117
QY 361  ACAGGGAAGTGACATTGGCCGAGGACCCCGGCC 396
DB 1118  ACAGGGAAGTGACATTGGCCGAGGACCCCGGCC 1153

RESULT 10
US-09-759-143-822
; Sequence 822, Application US/09759143
; Patent No. US2002022248A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
;           Dillon, Davin C.
;           Mitchell, Jennifer L.
;           Harlocker, Susan L.
;           Jiang, Yuqi
;           Henderson, Robert A.
;           Kalos, Michael D.
;           Fanger, Gary R.
;           Retter, Marc W.
;           Stolk, John A.
;           Day, Craig H.
;           Vedrick, Thomas S.
;           Carter, Darrick
;           Li, Samuel
;           Wang, Aijun
```

APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Hepler, William  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
FILE REFERENCE: 210121.427C23  
CURRENT APPLICATION NUMBER: US/09/759,143  
CURRENT FILING DATE: 2001-01-12  
NUMBER OF SEQ ID NOS: 934  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 822  
LENGTH: 675  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-759-143-822

Query Match 89.4%; Score 354; DB 9; Length 675;  
Best Local Similarity 95.7%; Pred. No. 7.le-92;  
Matches 379; Conservative 0; Mismatches 5; Indels 12; Gaps 1;

Qy 1 ACGGCCGCTCCGATAACTTCCAGCTGTCCAGGGTGGCGAGGATTCCGATTCCGATC 60  
Db 22 ACGGCCGCTCCGATAACTTCCAGCTGTCCAGGGTGGCGAGGATTCCGATTCCGATC 81  
Qy 61 GGGCAGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGTCAACCCACCGTTTCAT 120  
Db 82 GGGCAGCGATGGCGATCGCGGGCCAGATC-----AAGCTTCCACCGTTTCAT 129  
Qy 121 ATCCGGGCTACCGCTTCTCGGCTTGGGTGTTCGACAAACAAACGCGGCGACGA 180  
Db 130 ATCCGGGCTACCGCTTCTCGGCTTGGGTGTTCGACAAACAAACGCGGCGACGA 189  
Qy 181 GTCCAAACGCTGTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGACGTG 240  
Db 190 GTCCAAACGCTGTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGACGTG 249  
Qy 241 ATACCGGGTCGACGGGCTCCGATCAATCGGCCACCGGATGGCGGCGCTTAAAC 300  
Db 250 ATACCGGGTCGACGGGCTCCGATCAATCGGCCACCGGATGGCGGCGCTTAAAC 309  
Qy 301 GGGCATCATCCCGTGACGTCTCGGTGACCTGGCAAAACAAAGTGGGGGGGCGACGCGT 360  
Db 310 GGGCATCATCCCGTGACGTCTCGGTGACCTGGCAAAACAAAGTGGGGGGGCGACGCGT 369  
Qy 361 ACAGGGAACGTGACATTTGGCGGAGGACCCCGGCC 396  
Db 370 ACAGGGAACGTGACATTTGGCGGAGGACCCCGGCC 405

## RESULT 11

US-09-780-669-822  
Sequence 822, Application US/09780669  
Patent No. US2002005197A1  
GENERAL INFORMATION:  
APPLICANT: Xu, Jiangchun  
APPLICANT: Dillon, Devin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Jiang, Yuqi  
APPLICANT: Henderson, Robert A.  
APPLICANT: Kalos, Michael D.  
APPLICANT: Fanger, Gary R.  
APPLICANT: Retter, Marc W.  
APPLICANT: Stolk, John A.  
APPLICANT: Day, Craig H.  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Carter, Barrick  
APPLICANT: Li, Samuel  
APPLICANT: Wang, Aijun  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Hepler, William  
APPLICANT: Hural, John  
APPLICANT: McNeill, Patricia D.  
APPLICANT: Houghton, Raymond L.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
FILE REFERENCE: 210121.427C24  
CURRENT APPLICATION NUMBER: US/09/780.669  
CURRENT FILING DATE: 2001-02-09  
NUMBER OF SEQ ID NOS: 943  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 822  
LENGTH: 675  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-780-669-822

Query Match 89.4%; Score 354; DB 9; Length 675;  
Best Local Similarity 95.7%; Pred. No. 7.le-92;  
Matches 379; Conservative 0; Mismatches 5; Indels 12; Gaps 1;

Qy 1 ACGGCCGCTCCGATAACTTCCAGCTGTCCAGGGTGGCGAGGATTCCGATTCCGATC 60  
Db 22 ACGGCCGCTCCGATAACTTCCAGCTGTCCAGGGTGGCGAGGATTCCGATTCCGATC 81  
Qy 61 GGGCAGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGTCAACCCACCGTTTCAT 120  
Db 82 GGGCAGCGATGGCGATCGCGGGCCAGATC-----AAGCTTCCACCGTTTCAT 129  
Qy 121 ATCCGGGCTACCGCTTCTCGGCTTGGGTGTTCGACAAACAAACGCGGCGACGA 180  
Db 130 ATCCGGGCTACCGCTTCTCGGCTTGGGTGTTCGACAAACAAACGCGGCGACGA 189  
Qy 181 GTCCAAACGCTGTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGACGTG 240  
Db 190 GTCCAAACGCTGTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGACGTG 249  
Qy 241 ATACCGGGTCGACGGGCTCCGATCAATCGGCCACCGGATGGCGGCGCTTAAAC 300  
Db 250 ATACCGGGTCGACGGGCTCCGATCAATCGGCCACCGGATGGCGGCGCTTAAAC 309  
Qy 301 GGGCATCATCCCGTGACGTCTCGGTGACCTGGCAAAACAAAGTGGGGGGGCGACGCGT 360  
Db 310 GGGCATCATCCCGTGACGTCTCGGTGACCTGGCAAAACAAAGTGGGGGGGCGACGCGT 369  
Qy 361 ACAGGGAACGTGACATTTGGCGGAGGACCCCGGCC 396  
Db 370 ACAGGGAACGTGACATTTGGCGGAGGACCCCGGCC 405

## RESULT 12

US-09-822-827-822  
Sequence 822, Application US/09822827  
Patent No. US20020081680A1  
GENERAL INFORMATION:  
APPLICANT: Xu, Jiangchun  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
FILE REFERENCE: 210121.534C1  
CURRENT APPLICATION NUMBER: US/09/822.827  
CURRENT FILING DATE: 2001-03-28  
NUMBER OF SEQ ID NOS: 982  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 822  
LENGTH: 675  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-822-827-822

Query Match 89.4%; Score 354; DB 9; Length 675;  
Best Local Similarity 95.7%; Pred. No. 7.le-92;  
Matches 379; Conservative 0; Mismatches 5; Indels 12; Gaps 1;

Qy 1 ACGGCCGCTCCGATAACTTCCAGCTGTCCAGGGTGGCGAGGATTCCGATTCCGATC 60  
Db 22 ACGGCCGCTCCGATAACTTCCAGCTGTCCAGGGTGGCGAGGATTCCGATTCCGATC 81

QY 61 GGGCAGGCGATGGCGATCGGGCCAGATCCGATCGGGTGGGGGTGTCACCCACCGTTTCAT 120  
Db 82 GGGCAGGCGATGGCGATCGGGCCAGATC-----AAGCTTCCACCGTTTCAT 129  
QY 121 ATCGGGCCTACCGCTTCTCGGCTTGGGTGTTGTCGACAAACGCGCAACGGCGCACGA 180  
Db 130 ATCGGGCCTACCGCTTCTCGGCTTGGGTGTTGTCGACAAACGCGCAACGGCGCACGA 189  
QY 181 GTCCAAACGGTGGTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGCACGTG 240  
Db 190 GTCCAAACGGTGGTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGCACGTG 249  
QY 241 ATACCGCGGTGCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300  
Db 250 ATACCGCGGTGCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309  
QY 301 GGGCATCATCCCGGTGACGTGACGTGACGTGACGTGACGTGACGTGACGTGACGTGACGT 360  
Db 310 GGGCATCATCCCGGTGACGTGACGTGACGTGACGTGACGTGACGTGACGTGACGTGACGT 369  
QY 361 ACAGGGAACGTGACATTTGGCCGAGGAGACCCCGGGCC 396  
Db 370 ACAGGGAACGTGACATTTGGCCGAGGAGACCCCGGGCC 405

## RESULT 13

US-09-895-793-822

; Sequence 822, Application US/09895793

; Publication No. US20020192763A1

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; APPLICANT: Dillon, Davin C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Harlocker, Susan L.

; APPLICANT: Jiang, Yuqiu

; APPLICANT: Kalos, Michael D.

; APPLICANT: Retter, Marc W.

; APPLICANT: Stolk, John A.

; APPLICANT: Day, Craig H.

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Carter, Darrick

; APPLICANT: Li, Samuel X.

; APPLICANT: Wang, Aljun

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Hepler, William T.

; APPLICANT: Henderson, Robert A.

; APPLICANT: Hural, John

; APPLICANT: McNeill, Patricia D.

; APPLICANT: Houghton, Raymond L.

; APPLICANT: Vinals de Bassols, Carlota

; APPLICANT: Foy, Teresa

; APPLICANT: Fanger, Gary R.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; FILE REFERENCE: 210121.534C2

; CURRENT APPLICATION NUMBER: US/09/895,793

; CURRENT FILING DATE: 2001-06-29

; NUMBER OF SEQ ID NOS: 982

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 822

; LENGTH: 675

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-895-793-822

Query Match 89.4%; Score 354; DB 10; Length 675;

Best Local Similarity 95.7%; Pred. No. 7.1e-92;

Matches 379; Conservative 0; Mismatches 5; Indels 12; Gaps 1;

QY 1 ACGGCGCGGTCCGATAACTTCCAGTGTCCCGAGGTGGCGAGGGATTCCCGATTCGGATC 60  
Db 22 ACGGCGCGGTCCGATAACTTCCAGTGTCCCGAGGTGGCGAGGGATTCCCGATTCGGATC 81

QY 61 GGGCAGGCGATGGCGATCGGGCCAGATCCGATCGGGTGGGGGTGTCACCCACCGTTTCAT 120  
Db 82 GGGCAGGCGATGGCGATCGGGCCAGATC-----AAGCTTCCACCGTTTCAT 129  
QY 121 ATCGGGCCTACCGCTTCTCGGCTTGGGTGTTGTCGACAAACGCGCAACGGCGCACGA 180  
Db 130 ATCGGGCCTACCGCTTCTCGGCTTGGGTGTTGTCGACAAACGCGCAACGGCGCACGA 189  
QY 181 GTCCAAACGGTGGTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGCACGTG 240  
Db 190 GTCCAAACGGTGGTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGCACGTG 249  
QY 241 ATACCGCGGTGCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300  
Db 250 ATACCGCGGTGCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309  
QY 301 GGGCATCATCCCGGTGACGTGACGTGACGTGACGTGACGTGACGTGACGTGACGTGACGT 360  
Db 310 GGGCATCATCCCGGTGACGTGACGTGACGTGACGTGACGTGACGTGACGTGACGTGACGT 369  
QY 361 ACAGGGAACGTGACATTTGGCCGAGGAGACCCCGGGCC 396  
Db 370 ACAGGGAACGTGACATTTGGCCGAGGAGACCCCGGGCC 405

## RESULT 14

US-09-895-814-822

; Sequence 822, Application US/09895814

; Publication No. US20020193296A1

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; APPLICANT: Dillon, Davin C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Harlocker, Susan L.

; APPLICANT: Jiang, Yuqiu

; APPLICANT: Kalos, Michael D.

; APPLICANT: Retter, Marc W.

; APPLICANT: Stolk, John A.

; APPLICANT: Day, Craig H.

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Carter, Darrick

; APPLICANT: Li, Samuel X.

; APPLICANT: Wang, Aljun

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Hepler, William T.

; APPLICANT: Henderson, Robert A.

; APPLICANT: Hural, John

; APPLICANT: McNeill, Patricia D.

; APPLICANT: Houghton, Raymond L.

; APPLICANT: Vinals de Bassols, Carlota

; APPLICANT: Foy, Teresa

; APPLICANT: Fanger, Gary R.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; FILE REFERENCE: 210121.427C26

; CURRENT APPLICATION NUMBER: US/09/895,814

; CURRENT FILING DATE: 2001-06-29

; NUMBER OF SEQ ID NOS: 990

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 822

; LENGTH: 675

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-895-814-822

Query Match 89.4%; Score 354; DB 10; Length 675;

Best Local Similarity 95.7%; Pred. No. 7.1e-92;

Matches 379; Conservative 0; Mismatches 5; Indels 12; Gaps 1;

QY 1 ACGGCGCGGTCCGATAACTTCCAGTGTCCCGAGGTGGCGAGGGATTCCCGATTCGGATC 60  
Db 22 ACGGCGCGGTCCGATAACTTCCAGTGTCCCGAGGTGGCGAGGGATTCCCGATTCGGATC 81

```

Qy 61 GGGCAGCGATGGGATCGCGGGCCAGATCCGATCGGGTGGGGGTCAACCACCGTTTCAT 120
Db 82 GGGCAGCGATGGGATCGCGGGCCAGATCCGATCGGGTGGGGGTCAACCACCGTTTCAT 129
Qy 121 ATCGGGCTACCGCTTCCCTGGCTTGGGTGTTGTCGACAAACGCGGACGCGCAGCA 180
Db 130 ATCGGGCTACCGCTTCCCTGGCTTGGGTGTTGTCGACAAACGCGGACGCGCAGCA 189
Qy 181 GTCCAAACGCGTGGTTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGCAGTG 240
Db 190 GTCCAAACGCGTGGTTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGCAGTG 249
Qy 241 ATACCCCGGTGACGCGCGTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAAC 300
Db 250 ATACCCCGGTGACGCGCGTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAAC 309
Qy 301 GGGCATCATCCCGGTGACGCGCGTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCT 360
Db 310 GGGCATCATCCCGGTGACGCGCGTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCT 369
Qy 361 ACAGGGAACGTGACATTTGGCGGAGGACCCCGGCC 396
Db 370 ACAGGGAACGTGACATTTGGCGGAGGACCCCGGCC 405

```

RESULT 15

```

US-10-144-678A-822
; Sequence 822, Application US/10144678A
; Publication No. US20030157089A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jlangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Hepler, William T.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals y de Bassols, Carlota
; APPLICANT: Foy, Teresa M.
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Deng, Ta
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C28
; CURRENT APPLICATION NUMBER: US/10/144.678A
; NUMBER OF SEQ ID NOS: 1033
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 822
; LENGTH: 675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-144-678A-822

```

```

Query Match 89.4%; Score 354; DB 12; Length 675;
Best Local Similarity 95.7%; Pred. No. 7.le-92;
Matches 379; Conservative 0; Mismatches 5; Indels 12; Gaps 1;

Qy 1 ACAGCGCGTCCGATACCTTCCAGCTGTCAGGGTGGGAGGATTCGCCATTCGGATC 60
|

```

```

Db 22 ACAGCGCGTCCGATACCTTCCAGCTGTCAGGGTGGGAGGATTCGCCATTCGGATC 81
Qy 61 GGGCAGCGATGGGATCGCGGGCCAGATCCGATCGGGTGGGGGTCAACCACCGTTTCAT 120
Db 82 GGGCAGCGATGGGATCGCGGGCCAGATCCGATCGGGTGGGGGTCAACCACCGTTTCAT 129
Qy 121 ATCGGGCTACCGCTTCCCTGGCTTGGGTGTTGTCGACAAACGCGGACGCGCAGCA 180
Db 130 ATCGGGCTACCGCTTCCCTGGCTTGGGTGTTGTCGACAAACGCGGACGCGCAGCA 189
Qy 181 GTCCAAACGCGTGGTTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGCAGTG 240
Db 190 GTCCAAACGCGTGGTTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGCAGTG 249
Qy 241 ATACCCCGGTGACGCGCGTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAAC 300
Db 250 ATACCCCGGTGACGCGCGTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAAC 309
Qy 301 GGGCATCATCCCGGTGACGCGCGTCCGATCAACTCGGCCACCGCGATGGCGGACGCGGT 360
Db 310 GGGCATCATCCCGGTGACGCGCGTCCGATCAACTCGGCCACCGCGATGGCGGACGCGGT 369
Qy 361 ACAGGGAACGTGACATTTGGCGGAGGACCCCGGCC 396
Db 370 ACAGGGAACGTGACATTTGGCGGAGGACCCCGGCC 405

```

Search completed: September 5, 2003, 08:54:31  
Job time : 1536 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2003, 07:20:04 ; Search time 2067 Seconds  
(without alignments)  
4556.300 Million cell updates/sec

Title: US-09-684-215A-3  
Perfect score: 396  
Sequence: 1 acggcggtccgataact.....tggccgaggagcccccgcc 396

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*

1: em\_estba:\*\*  
2: em\_esthum:\*\*  
3: em\_estin:\*\*  
4: em\_estmu:\*\*  
5: em\_estor:\*\*  
6: em\_estpl:\*\*  
7: em\_estro:\*\*  
8: em\_htc:\*\*  
9: gb\_est1:\*\*  
10: gb\_est2:\*\*  
11: gb\_hc:\*\*  
12: gb\_est3:\*\*  
13: gb\_est4:\*\*  
14: gb\_est5:\*\*  
15: em\_estfun:\*\*  
16: em\_estom:\*\*  
17: em\_gss\_hum:\*\*  
18: em\_gss\_inv:\*\*  
19: em\_gss\_pln:\*\*  
20: em\_gss\_vrt:\*\*  
21: em\_gss\_fun:\*\*  
22: em\_gss\_mam:\*\*  
23: em\_gss\_mus:\*\*  
24: em\_gss\_pro:\*\*  
25: em\_gss\_rod:\*\*  
26: em\_gss\_phg:\*\*  
27: em\_gss\_vrl:\*\*  
28: gb\_gss1:\*\*  
29: gb\_gss2:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	65.2	16.5	289	29	U82114
2	55.6	14.0	603	28	AZ934428 BJ_Ba000
3	50.4	12.7	925	29	CNS0091P
4	49.6	12.5	925	29	CNS0091P

5	49.2	12.4	561	28	BH020987
6	48.6	12.3	634	10	BE361689
7	47.8	12.1	831	29	CC334434
8	47.6	12.0	289	29	P864L
9	47.4	12.0	935	29	CNS006XK
10	47	11.9	1009	29	CNS010EW
11	46.4	11.7	515	13	AX424977
12	45.8	11.6	472	14	CB620050
13	45.6	11.5	515	13	AX424977
14	45.4	11.5	459	29	CC349367
15	45.4	11.5	830	29	CC331586
16	45.4	11.4	538	10	BE215641
17	44.8	11.3	502	6	AU191977
18	44.6	11.3	839	29	CNS004NB
19	44.6	11.3	1200	13	AX456467
20	44.6	11.3	1201	13	AX360624
21	44.4	11.2	496	12	BJ307565
22	44.4	11.2	504	28	BH190076
23	44.4	11.2	590	28	BH189840
24	44.4	11.2	632	28	BH189865
25	44.4	11.2	932	29	CNS00720
26	44.2	11.2	844	29	CNS0052P
27	44	11.1	493	10	BE355002
28	44	11.1	536	10	EG412839
29	44	11.1	575	14	CD222185
30	44	11.1	587	14	CD229801
31	44	11.1	618	14	CD221131
32	44	11.1	644	14	CD226234
33	43.8	11.1	527	14	CB636890
34	43.8	11.1	538	14	CB636749
35	43.8	11.1	767	14	CB676645
36	42.6	10.8	442	12	BM523115
37	42.6	10.8	498	10	BG411096
38	42.4	10.7	638	14	CD220173
39	42.4	10.7	959	29	BZ549048
40	42.2	10.7	578	14	CB636482
41	42.2	10.7	1201	13	AX360624
42	42.2	10.7	1201	13	AX460099
43	42	10.6	755	10	BE705003
44	42	10.6	1004	13	BU772046
45	42	10.6	1033	29	BZ561390

## ALIGNMENTS

RESULT 1  
U82114  
LOCUS  
DEFINITION U82114 ordered cosmid library Mycobacterium leprae genomic clone  
cosmid L-373; contig 64, genomic survey sequence.  
ACCESSION U82114  
VERSION U82114.1 GI:3647212  
KEYWORDS GSS.  
SOURCE Mycobacterium leprae  
ORGANISM Mycobacterium leprae  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
REFERENCE 1 (bases 1 to 289)  
AUTHORS Silbag,F.S., Cho,S.N., Cole,S.T. and Brennan,P.J.  
TITLE Characterization of a 34-kilodalton protein of Mycobacterium leprae that is homologous to the immunodominant 34-kilodalton antigen of Mycobacterium paratuberculosis  
JOURNAL Infect. Immun. 66 (11), 5576-5579 (1998)  
MEDLINE 99003163  
PUBMED 9784577  
COMMENT Contact: Silbag FS  
Microbiology  
Colorado State University  
Fort Collins, CO 80523, USA  
Eiglmeyer,K., Honore,N., Woods,S.A., Caudron,B. and Cole,S.T. Use of an ordered cosmid library to deduce the genomic organization of Mycobacterium leprae. Mol. Microbiol. 7 (2), 197-206 (1993)

289 bp DNA linear GSS 16-FEB-2001  
Mycobacterium leprae genomic clone

Query Match	14.08;	Score 55.6;	DB 28;	Length 603;
Best Local Similarity	54.4%;	Pred. No. 0.00099;		
Matches	112;	Conservative 0;	Mismatches 94;	Indels 0; Gaps 0
QY	140	TCGGCTTGGTGTTCGACAAACGCGACGCGACGAGTCCAAACGCGTGTTCGGGA	199	
Db	363	TCGCCGACACCCCTCGGCATGATGAAGGCCAAGGGCGCTGGTGGCGGAGCCGACGGCGA	422	
QY	200	GCCTCTCGGGCGCAAGTCTCGGATCTCCACCGCGACGTCATCACCGCGGTTCGACGCGC	259	
Db	423	ACGGTCTCGGGCGCAAGGCCGCGATCGAGTCCGCGACGTCATCACCTCGTCAAGCGC	482	
QY	260	CTCCGATCAACTCGGCCACCGCGATGCGCGACGCGCTTAACGGGCATCATCCCGGTGACG	319	
Db	483	AATCCGTCGAAGGACGCGCGGAGTCGCCGCGACCATCGCGGCATGCGCGCGGTGCGA	542	
QY	320	TCATCTCGGTGACCTGSCAAACCAAG	345	
Db	543	TCCTGAAGCTTAAGTGTCTCACAAG	568	
RESULT 3				
CNS0091P				
LOCUS	CNS0091P 925 bp DNA linear GSS 03-JUN-1999			
DEFINITION	Drosophila melanogaster genome survey sequence TET3 end of BAC #			
ACCESSION	BACR19D16 of RPCI-98 library from Drosophila melanogaster (fruit			
VERSION	fly), genomic survey sequence.			
KEYWORDS	AL053013			
SOURCE	AL053013.1 GI:4934461			
ORGANISM	GSS.			
REFERENCE	Drosophila melanogaster (fruit fly)			
AUTHORS	Drosophila melanogaster			
TITLE	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
JOURNAL	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
COMMENT	Ephydroidea; Drosophilidae; Drosophila.			
	1 (bases 1 to 925)			
	Genoscope.			
	Direct Submission			
	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :			
	BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr			
	- Web : www.genoscope.cns.fr)			
	Determination of this BAC-end sequence was carried out as part of a			
	collaboration with the Berkeley Drosophila Genome Project (BDGP).			
	The BDGP is constructing a physical map of the Drosophila			
	melanogaster genome using these BACs. For further information			
	please see <a href="http://www.fruitfly.org">http://www.fruitfly.org</a> The BDGP Drosophila			
	melanogaster BAC library was prepared by Kazutoyo Osoegawa and			
	Aaron Mammoler in Pletier de Jong's laboratory in the Department of			
	Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,			
	NY. The library is named RPCI-98 and was constructed by partial			
	EcoRI digestion of Drosophila DNA provided by the BDGP from the			
	isogenic strain Y2; cn bw sp, the same strain used for the BDGP's			
	p1 and EST libraries. A more detailed description of the library			
	and how to order individual BAC clones, the entire library, or			
	filters for hybridization from the BACPAC Resource Center can be			
	found at <a href="http://bacpac.med.buffalo.edu/drosophila_bac.htm">http://bacpac.med.buffalo.edu/drosophila_bac.htm</a> .			
FEATURES	Location/Qualifiers			
source	1..925			
	/organism="Drosophila melanogaster"			
	/mol_type="genomic DNA"			
	/db_xref="taxon:7227"			
	/clone="BACR19D16"			
	/clone_lib="RPCI-98"			
	/note="end : TET3"			
BASE COUNT	120 a	61 c	61 g	172 t
ORIGIN	511 others			
Query Match	12.7%;	Score 50.4;	DB 29;	Length 925;
Best Local Similarity	15.4%;	Pred. No. 0.02;		
Matches	53;	Conservative 158;	Mismatches 133;	Indels 1; Gaps 1;
Query Match	53	TTCGATCGGGCAGGCGATCGCATCGCGGGCCAGATCCGATCGGTGGGGGTCAACCCA	112	



```
Db 553 TTSSGSGYGGKSGSGBSCSCSSCSBSSSSCCBCCSCSSSYCCSSSSSSSKSS 612
Qy 113 CCGTTATATACGGCCCTACCGCTCTCCGCTTGGGTGTTCGACAAACGCAACG 172
Db 613 TSBSGCCSCSKSVCGTSCSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 672
Qy 173 GCGACAGTCCACACGGG-TGGTGGGAGCGCTCCGCGGCAAGTCTCGCATCTCC 231
Db 673 KTSASGSGWSAGGSGTSTSSSSSSSSSTSSSSSVSSGSKSTBSSGSSSSGSS 732
Qy 232 GCGACGTGATCACCAGGCTCGACGGCGCTCCGATCACTCGCCACCAGGCGGAC 291
Db 733 SSSTSSBBSCTSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 792
Qy 292 GCGCTTAACGGGATCATCCGCTGAGCTCATCTCGGTGACCTGGCAACCAAGTC 351
Db 793 GTSSSSDSTSCCCCYMCTCCSTYBMCYTSTSCGSSSSSSGKGVTCKCGCGG 852
Qy 352 GGCACGGTACAGGAACGTGACATTCGCGGAGGACCCCGCGCC 396
Db 853 TNGMBGTSSACSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 897

RESULT 4
CNS0091P/c
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACR19D16 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION
AL053013
VERSION
AL053013.1 GI:4934461
KEYWORDS
GSS.
SOURCE
Drosophila melanogaster (fruit fly)
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 925)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osogawa and
Aaron Mamoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES
Location/Qualifiers
1..925
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR19D16"
/clone_lib="RPCI-98"
/note="end : TET3"
BASE COUNT 120 a 61 c 61 g 172 t 511 others
ORIGIN
Query Match 12.5%; Score 49.6; DB 29; Length 925;
Best Local Similarity 12.8%; Pred. No. 0.031;
Matches 43; Conservative 157; Mismatches 136; Indels 0; Gaps 0;
```

```
Qy 60 CGGACGCGATCGCGGCGCCAGATCCGATCGGTGGGGGTTCACCCACCGTTCA 119
Db 900 NSSBSCSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 841
Qy 120 TATCGCGCTACCGCTTCTCGCTTGGTGTTCGACAAACGCAACGCGCGCACG 179
Db 840 BCMCMSSSSSSCCGSASARGVKVRASGAGRGKGGSGGASASHSSSSACBSSS 781
Qy 180 AGTCCAACGCGTGTGGGAGCGCTCCGCGCGCAAGTCTCGCATCTCCACCGGCG 239
Db 780 SASSSSSASRRSRGGAGGSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 721
Qy 240 GATCACCGCGCTACGCGCTCCGATCAACTCGGCCACCGCGCGACGCGCTTAA 299
Db 720 VASSMCSBSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 661
Qy 300 CGGCGATCATCCGCTGACATCTCGGTGACCTGCGCAACCAAGTCGGCGGCG 359
Db 660 CSSSSMSASSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 601
Qy 360 TACAGGGAACGTGACATTCGCGGAGGACCCCGCGC 395
Db 600 GRSSGSGGGVGGSGSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 565

RESULT 5
BH020987
LOCUS
DEFINITION
P864c.d_LEISHPAC-left.1 Leishmania major Friedlin Cosmid Genomic
Library Leishmania major genomic clone P864c, genomic survey
sequence.
ACCESSION
BH020987
VERSION
BH020987.1 GI:14202102
KEYWORDS
GSS.
SOURCE
Leishmania major
ORGANISM
Leishmania major
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
Leishmania.
REFERENCE
1 (bases 1 to 561)
Mylers, P.J., Vogt, C., Cawthra, J., Klacking, M., Marty, A., Mack, J.,
Munden, H., Nguyen, D., Robertson, L., Sisk, E., Fazelinia, G., Aggarwal
, G., Nelson, S., Seyler, A., Worthey, E. and Stuart, K.
Leishmania major Friedlin Cosmid End Sequences
Unpublished
Other GSSs: P864c.d_LEISHPAC-right.1
Contact: Myler PJ
Seattle Biomedical Research Institute
4 Nickerson Street, Seattle, WA 98109-1651, USA
Tel: 206 284-8846
Fax: 206 284-0313
Email: mylerpj@sbri.org
Seq primer: LEISHPAC-Left
Class: PAC end.
FEATURES
Location/Qualifiers
1..561
/organism="Leishmania major"
/mol_type="genomic DNA"
/strain="Friedlin"
/db_xref="taxon:5664"
/clone="P864c"
/lab_host="E. coli ED8767"
/clone_lib="Leishmania major Friedlin Cosmid Genomic
Library"
/note="Vector: cLHYG; Site.1: BamHI; Genomic DNA from
Leishmania major Friedlin was partially digested with
Sau3AI, size selected, and ligated with BamHI-digested
cLHYG cosmid vector DNA. 9216 clones were picked and
arrayed. Library construction is described in Ivens et
al., Genomics Research, 8:135-145 (1998). The cLHYG
vector (Acc. No. CVU59231) is described in Ryan et al.,
Gene, 131:145-150 (1993)."
BASE COUNT 108 a 173 c 182 g 98 t
ORIGIN
```

Db	268	ACGAGCTCGTCTGCTCTGCTCTGCTCTCGGAGTCGGCGCGGGAACAGCCCTCTCTGCTGCT	209
Qy	197	GGAGGCTCCGCGCGCAAGTCTCGGCACTCCACCGCGGAGCGTATCACCGCGGTGCAGC	256
Db	208	GCGCGCGCGCGCGCGCGCTACTGCTCCGCGGTGGCGGCTTCGCGTGGCTGGACTTGT	149
Qy	257	GGGCTCCGATGAATCTCGGCCACCGCGAGTGGCGGACGCGCTTAACGGGCATCATCCGGTG	316
Db	148	GCCCGCGAGCGCCTGGTACGACGGAAGGCTTCCGCGACACCGCGCACCGGAACCGGA	89
Qy	317	ACGTCAATCTCGGTGACCTG	335
Db	88	GCTGCTCTCGGCGCCAG	70

RESULT 7	CC334434	831 bp	DNA	linear	GSS 16-MAY-2003
LOCUS	CC334434.1	7,151	base pairs	linear	16-MAY-2003
DEFINITION	CC334434.1	7,151	base pairs	linear	16-MAY-2003
ACCESSION	CC334434	7,151	base pairs	linear	16-MAY-2003
VERSION	CC334434.1	7,151	base pairs	linear	16-MAY-2003
KEYWORDS	CC334434.1	7,151	base pairs	linear	16-MAY-2003
SOURCE	CC334434.1	7,151	base pairs	linear	16-MAY-2003
ORGANISM	CC334434.1	7,151	base pairs	linear	16-MAY-2003
REFERENCE	CC334434.1	7,151	base pairs	linear	16-MAY-2003
AUTHORS	CC334434.1	7,151	base pairs	linear	16-MAY-2003
TITLE	CC334434.1	7,151	base pairs	linear	16-MAY-2003
JOURNAL	CC334434.1	7,151	base pairs	linear	16-MAY-2003
COMMENT	CC334434.1	7,151	base pairs	linear	16-MAY-2003

```

Class: sheared ends.
Location/Qualifiers
1. .83p
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBma0365N21"
/clone_lib="ZM_0.7_1.5_KB"
/note="vector: pBCSK-; Site: 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"
BASE COUNT 149 a 280 c 290 g 112 t
ORIGIN
Query Match
Best Local Similarity 50.1%; Score 47.8; DB 29; Length 831;
Matches 109; Conservative 0; Mismatches 77; Indels 5; Gaps 1;
Qy 164 ACGGCAACGGCGACGAGTCCAGTCCAAAGCGGTGTCGGGAGCGCTCGCGGGCAAGCTTCGGCA 223
Db 777 ACGGGACAGCGCGGAAGCGCGGTTCGTGCTCCGGCAGGATCGGCGG-----CTGCCCG 723
Qy 224 TCTCCACCGCGGAGGTGATCACCGGGTTCGACGGCGCTCGGATCACTCGGCGACACGCGCA 283
Db 722 TCGCCACCCCTCGAGTTCGGCGCTCGACGCCGCCCGCGCTGCACGGCGCGCGCGCGC 663
Qy 284 TGCGGGACGGCGCTAACGGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAACCA 343
Db 662 GCGCCTTCGGGTCCGGGGCGGCGCTCCTCCGCGACGTCTCTCTCTCTCTCTCTCTCTCGCGTCCG 603
Qy 344 AGTCGGGCGGC 354

```

```

Db 602 CCCC GCCCGC 592

RESULT 8
P864L
LOCUS
DEFINITION Leishmania major Friedlin PAC P864 left end-sequence, genomic survey sequence.
ACCESSION AL390548
VERSION AL390548.1 GI:9501524
KEYWORDS GSS
SOURCE Leishmania major
ORGANISM Leishmania major
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
REFERENCE
AUTHORS Ivens,A.C., Lewis,S.M., Bagherzadeh,A., Zhang,L., Chan,H.M. and Smith,D.F.
1 (bases 1 to 289)
A physical map of the Leishmania major Friedlin genome
Genome Res. 8 (2), 135-145 (1998)
9477341
PUBMED 98146435
REFERENCE
AUTHORS Taylor,R.G., Huckle,E.E.J., Ivens,A.C., Rajandream,M.A. and Barrell,B.G.
2 (bases 1 to 289)
Direct Submission
Submitted (24-JUL-2000) Leishmania major Friedlin genome sequencing project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and alicat@sanger.ac.uk
COMMENT see http://www.ebi.ac.uk/parasites/leish.html
Details of Leishmania sequencing at the Sanger Centre are available at http://www.sanger.ac.uk/Projects/L_major/
The primer sequence can be obtained from alicat@sanger.ac.uk.
FEATURES
Location/Qualifiers
source
1..289
/organism="Leishmania major"
/mol_type="genomic DNA"
/strain="Friedlin"
/db_xref="taxon:5664"
/clone="PAC P864"
BASE COUNT 49 a 86 c 108 g 46 t
ORIGIN
Query Match 12.0%; Score 47.6; DB 29; Length 289;
Best Local Similarity 62.7%; Pred. NO. 0.084;
Matches 74; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
QY 159 CAACACGGGCAACGGCGACGATCCACGGCGTGTGCGGGAGCGCTCGGGCGGCAAGTCT 218
DB 24 CGAGAGCGGATCGCCGGCGCCTGTATTAGTCGAGGGCGCGTCCGGTCCCGGGTCCGGT 83
QY 219 CGGATCTCCACCGCGACGTGATCACCGCGGTGACGGCGCTCGATCAACTCGGCC 276
DB 84 CGGATTCACCCGGCGACGTGATCTCGCTTCAACGGCAGCGGATCAAGTCAGTC 141

RESULT 9
NS006XK
LOCUS
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC # BAC14N09 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL066051
VERSION AL066051.1 GI:4945019
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
AUTHORS 1 (bases 1 to 935)
Genoscope.

Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osogawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial ECORI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
Location/Qualifiers
source
1..935
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BAC14N09"
/clone_lib="RPCI-98"
/note="end : T7"
BASE COUNT 257 a 170 c 162 g 96 t 250 others
ORIGIN
Query Match 12.0%; Score 47.4; DB 29; Length 935;
Best Local Similarity 28.2%; Pred. NO. 0.11;
Matches 111; Conservative 89; Mismatches 193; Indels 0; Gaps 0;
QY 4 GCCCGCCGATACTCCAGCTGTCCAGGTGGGAGGATTCGCCATTCCGATCGG 63
DB 523 GCSTCCMTWSSSVSCCGTCCGTCCTCKMCSCTYGGCKGCGCCSTCSSSSCCS 582
QY 64 CAGCGATGCGGATCGCGGGCGAGATCGGATGGGGGTGTCACCCACCGTTCATATC 123
DB 583 BBSYSTCCTBCTKCSGCGSTGCTCGCGGGGCGCGCGCGCGGGCGGCGSS 642
QY 124 GGGCTTACCGCTTCTCGGCTTGGTGTTCGACACACAGGCAAGCGCGCAGGTC 183
DB 643 GSSGSSGSSGSSGSSGSSGSSGSSGSSGSSGSSGSSGSSGSSGSSGSSG 702
QY 184 CAACGCGTGTGCGGAGCGCTCCGCGCAAGTCTCGGCATCTCCACCGCGAGTGC 243
DB 703 CGSGCGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 762
QY 244 ACCGCGTGTGACGGCGTCCGATCAACTCGGCCACCGCGATGGCGGACGCTTAACGG 303
DB 763 MSMCGGCGGSCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGSS 822
QY 304 CATCATCCGCTGAGTGCATCTCGGTGACCTGGCAACCAAGTGGGCGGCGCGGTACA 363
DB 823 GSCGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGSS 882
QY 364 GGAACGCTGACATTGGCGGAGGACCCCGCGCC 396
DB 883 GSGSGGCGGSGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 915

RESULT 10
NS010EW/C
LOCUS
DEFINITION Drosophila melanogaster genome survey sequence SP6 end of BAC BACN03P19 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL098832
VERSION AL098832.1 GI:5610493
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)

```







GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 5, 2003, 07:52:39 ; Search time 2356.88 Seconds  
(without alignments)  
2291.192 Million cell updates/sec

Title: US-09-684-215A-4  
Perfect score: 675  
Sequence: 1 TAASDNFQLSQGGQGAIPA.....QTKSGGTRGNVTLAGPPA 132

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlp  
-O=/cgn2\_1/USPTO.spool\_p/US09684215/runat\_05092003\_072202\_559/app\_query.fasta\_1.853  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=ptc -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORW=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09684215.ecgn\_1.1.3962 @runat\_05092003\_072202\_559 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_ov:\*  
5: gb\_pat:\*  
6: gb\_ph:\*  
7: gb\_pl:\*  
8: gb\_pr:\*  
9: gb\_ro:\*  
10: gb\_ro:\*  
11: gb\_ro:\*  
12: gb\_sy:\*  
13: gb\_sy:\*  
14: gb\_sy:\*  
15: gb\_sy:\*  
16: gb\_sy:\*  
17: gb\_sy:\*  
18: gb\_sy:\*  
19: gb\_sy:\*  
20: gb\_sy:\*  
21: gb\_sy:\*  
22: gb\_sy:\*  
23: gb\_sy:\*  
24: gb\_sy:\*  
25: gb\_sy:\*  
26: gb\_sy:\*  
27: gb\_sy:\*  
28: gb\_sy:\*

29: em\_vi:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pin:\*  
35: em\_htg\_rod:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	675	100.0	1068	6	AX005788 Sequence
2	675	100.0	1143	6	AX005790 Sequence
3	675	100.0	2287	6	AR303127 Sequence
4	675	100.0	11700	1	MTC1418B Mycobacteri
5	675	100.0	14029	1	AE006925 Mycobacte
6	675	100.0	343050	1	BX248334 Mycobacte
7	670	99.3	447	6	AR169152 Sequence
8	670	99.3	447	6	AR182442 Sequence
9	670	99.3	447	6	AR194825 Sequence
10	670	99.3	447	6	AR233097 Sequence
11	670	99.3	447	6	AX429596 Sequence
12	670	99.3	447	6	BD006325 Compounds
13	670	99.3	447	6	BD006445 Compounds
14	670	99.3	447	6	BD069285 Compounds
15	670	99.3	1872	6	AR169165 Sequence
16	670	99.3	1872	6	AR182455 Sequence
17	670	99.3	1872	6	AR194838 Sequence
18	670	99.3	1872	6	AR233110 Sequence
19	670	99.3	1872	6	AX429609 Sequence
20	670	99.3	1872	6	BD006338 Compounds
21	670	99.3	1872	6	BD006458 Compounds
22	670	99.3	1872	6	BD069298 Compounds
23	632	93.6	675	6	AX201049 Sequence
24	632	93.6	675	6	AX267848 Sequence
25	632	93.6	822	6	AR277645 Sequence
26	632	93.6	822	6	AR277645 Sequence
27	632	93.6	822	6	AX369152 Sequence
28	632	93.6	894	6	AX351489 Sequence
29	632	93.6	900	6	AR220690 Sequence
30	632	93.6	900	6	AX365960 Sequence
31	632	93.6	915	6	AR261281 Sequence
32	632	93.6	915	6	AX201061 Sequence
33	632	93.6	915	6	AX267860 Sequence
34	632	93.6	945	6	AR277644 Sequence
35	632	93.6	945	6	AX369151 Sequence
36	632	93.6	1012	6	AR220689 Sequence
37	632	93.6	1012	6	AX365958 Sequence
38	632	93.6	1155	6	AX369165 Sequence
39	632	93.6	1203	6	AX201078 Sequence
40	632	93.6	1203	6	AX267877 Sequence
41	632	93.6	1464	6	AR229410 Sequence
42	632	93.6	1464	6	AX156105 Sequence
43	632	93.6	1464	6	AX361955 Sequence
44	632	93.6	1557	6	AR229398 Sequence
45	632	93.6	1557	6	AX156089 Sequence

ALIGNMENTS

```
AX005788
LOCUS AX005788 1068 bp DNA linear PAT 24-AUG-2000
DEFINITION Sequence 907 from Patent WO9909186.
ACCESSION AX005788
VERSION AX005788.1 GI:9928795
KEYWORDS Mycobacterium tuberculosis
SOURCE Mycobacterium tuberculosis
ORGANISM Mycobacterium tuberculosis
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
REFERENCE
AUTHORS Portnoi, D. and Guigueno, A.
TITLE Polypeptide nucleic sequences exported from mycobacteria, vectors
comprising same and uses for diagnosing and preventing tuberculosis
JOURNAL Patent: WO 9909186-A 907 25-FEB-1999;
PORTNOI DENIS (FR); GUIGUENO AGNES (FR)
FEATURES
source Location/Qualifiers
1..1068
/organism="Mycobacterium tuberculosis"
/mol_type="genomic DNA"
/db_xref="taxon:1773"
1..1068
/contig="50D"
/codon_start=1
/transl_table=11
/protein_id="CAC05178.1"
/db_xref="GI:9928796"
/translation="MNSRRRLRWLLSLVLAAGLGLATAPAAQAPALSDQRFAD
FPALPLDPSAMVAQVGPQVNTKLGYNNAVGAGTGIVDPNGVVLNNHVIAGATD
INAFVSGGTYGVDPVGTDRQDVAVLQLRAGGLPSAAIGGGVAVGEPVVMGNSG
GOGGTPRAVPAGVAGLQTVQASDLTAETLNGLIQFDAAIQPDGSGGPPVNLGQ
VNGMTAASDNFQLSQGGGFAIPICQAMATAGQIRSGGSPTHVIGTAFGLGLVVD
NNGCARQVRVVGSAAPASLGISTGDTVITAVDGAIPNSATAMADALNGHHHPGDVIST
WQTKSGGTRTGNVTLAEGPPA"
BASE COUNT 169 a 352 c 375 g 172 t
ORIGIN
Alignment Scores:
Pred. No.: 3.25e-42 Length: 1068
Score: 675.00 Matches: 132
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-684-215A-4 (1-132) x AX005788 (1-1068)
QY 1 ThrAlaAlaSerAspAsnPhcInLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
DB 670 ACGGCGCGGTCGCGATTAACCTCCAGCTGTCAGGAGGCGGAGGATTCGCCATTCGATC 729
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
DB 730 GGGCAGGCGATGGCATCCGCGGCGCAGATCCATCGGTGGGGGTCCACCCACCGTTTAT 789
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAlaAspAsnGlyAlaArg 60
DB 790 ATCGGCGCTACCGCTCTCTCGGCTTGGGTGTGTCGACAAACAGCGGCGGCGACGA 849
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
DB 850 GTCCAAACGGTGGTGGGAGCGCTCCGCGGCAAGCTCCGGCATCTCCACCGCGGCGACGTG 909
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
DB 910 ATCAGCGCGGTGCGCGCGCTCCGATCACTACCGCGGCGGCGGATCGCGGCGGCTTAAC 969
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
DB 970 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGCAAAACCAAGTCGCGGCGGCGACGCT 1029
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
```

```
|||||
DB 1030 ACAGGGAACGTGACATTGGCCGAGGACCCCGGCC 1065
RESULT 2
AX005790
LOCUS AX005790 1143 bp DNA linear PAT 24-AUG-2000
DEFINITION Sequence 909 from Patent WO9909186.
ACCESSION AX005790
VERSION AX005790.1 GI:9928797
KEYWORDS Mycobacterium tuberculosis
SOURCE Mycobacterium tuberculosis
ORGANISM Mycobacterium tuberculosis
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
REFERENCE
AUTHORS Portnoi, D. and Guigueno, A.
TITLE Polypeptide nucleic sequences exported from mycobacteria, vectors
comprising same and uses for diagnosing and preventing tuberculosis
JOURNAL Patent: WO 9909186-A 909 25-FEB-1999;
PORTNOI DENIS (FR); GUIGUENO AGNES (FR)
FEATURES
source Location/Qualifiers
1..1143
/organism="Mycobacterium tuberculosis"
/mol_type="genomic DNA"
/db_xref="taxon:1773"
<4..1143
/contig="SEQ ID NO 50F"
/codon_start=1
/transl_table=11
/protein_id="CAC05179.1"
/db_xref="GI:9928798"
/translation="KNPAARTLKAGTISDYSYDPTGGYGMNSRRRLRWLLSLVLA
VGLGLATAQAPALSDQRFADFPALPLDPSAMVAQVGPQVNTKLGYNNAVGA
GTGIVDPNGVVLNNHVIAGATINAFVSGGTYGVDPVVGVDRTQDVAVLQLRAG
GLPSAAIGGGVAVGEPVVMGNSGCGGTTPRAVPGRVVAGLQTVQASDLTAETL
GLIQFDAAIQPDGSGGPPVNLGQVNGMTAASDNFQLSQGGGFAIPICQAMATAG
IRSGGSPTHVIGTAFGLGVNDNNGCARVQVRVVGSAAPASLGISTGDTVITAVDGA
PINSATAMADALNGHHHPGDVISTWQTKSGGTRTGNVTLAEGPPA"
BASE COUNT 189 a 373 c 395 g 186 t
ORIGIN
Alignment Scores:
Pred. No.: 3.48e-42 Length: 1143
Score: 675.00 Matches: 132
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-684-215A-4 (1-132) x AX005790 (1-1143)
QY 1 ThrAlaAlaSerAspAsnPhcInLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
DB 745 ACGGCGCGGTCGCGATTAACCTCCAGCTGTCAGGAGGCGGAGGATTCGCCATTCGATC 804
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
DB 805 GGGCAGGCGATGGCATCCGCGGCGCAGATCCATCGGTGGGGGTCCACCCACCGTTTAT 864
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAlaAspAsnGlyAlaArg 60
DB 865 ATCGGCGCTACCGCTCTCTCGGCTTGGGTGTGTCGACAAACAGCGGCGGCGACGA 924
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
DB 925 GTCCAAACGGTGGTGGGAGCGCTCCGCGGCAAGCTCCGGCATCTCCACCGCGGCGACGTG 984
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
DB 985 ATCAGCGCGGTGCGCGGCGCTCCGATCACTACCGCGGCGGCGGATCGCGGCGGCTTAAC 1044
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
```



```

|||||
1045 GGGCATCATCCGGTACGTCATCTCGGTACCTGGCAACCAAGTGGCGGCGACGCGT 1104
|||||
121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
|||||
1105 ACAGGGAACGTGACATTGGCCGAGGACCCCGGCC 1140

RESULT 3
AR303127
LOCUS AR303127 2287 bp DNA linear PAT 12-JUN-2003
DEFINITION Sequence 1 from patent US 6544522.
ACCESSION AR303127
VERSION AR303127.1 GI:31691855
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2287)
AUTHORS Skeiky,Y., Alderson,M. and Campos-Reto,A.
TITLE Fusion proteins of mycobacterium tuberculosis antigens and their
uses
JOURNAL Patent: US 6544522-A 1 08-APR-2003;
FEATURES
    source Location/Qualifiers
BASE COUNT 372 a 717 c 814 g 381 t 3 others
ORIGIN
1..2287
/organism="unknown"

Alignment Scores:
Pred. No.: 6,93e-42 Length: 2287
Score: 675.00 Matches: 132
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-684-215A-4 (1-132) x AR303127 (1-2287)

Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
|||||
Db 63 ACGCGCGCTCCGATACCTTCCAGCTGTCCAGGGTGGGAGGATCGGCATTCGGATC 122
|||||
Qy 21 GlyClnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
|||||
Db 123 GGGCAGCGCATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGTCCACCACCGTTTCAT 182
|||||
Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAlaArg 60
|||||
Db 183 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAAACACGCGGCGCACGA 242
|||||
Qy 61 ValClnArgValValGlySerAlaProAlaAlaSerLeuGlyTleSerThrGlyAspVal 80
|||||
Db 243 GTCCACGCGTGTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGAGTG 302
|||||
Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
|||||
Db 303 ATCACCAGCGTGCAGCGCGCTCCGATCACTCGGCCACCGCATGGCGGAGCGGCTTAAC 362
|||||
Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpClnThrIlySerGlyGlyThrArg 120
|||||
Db 363 GGGCATCATCCGGTACGTCATCTCGGTACCTGGCAACCAAGTGGCGGCGGCGACGCGT 422
|||||
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
|||||
Db 423 ACAGGGAACGTGACATTGGCCGAGGACCCCGGCC 458

RESULT 4
MTCI418B
LOCUS MTCI418B 11700 bp DNA linear BCT 02-SEP-2002
DEFINITION Mycobacterium tuberculosis H37Rv complete genome; segment 7/162.
ACCESSION Z96071 AL123456
VERSION Z96071.1 GI:3242254
KEYWORDS

```

SOURCE  
ORGANISM

Mycobacterium tuberculosis H37Rv  
Mycobacterium tuberculosis H37Rv  
Bacteria: Actinobacteria: Actinobacteridae: Actinomycetales:  
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium  
tuberculosis complex.

REFERENCE  
AUTHORS

1  
Cole,S.T., Brosch,R., Parkhill,J., Garnier,T., Churcher,C.,  
Harris,D., Gordon,S.V., Eiglmeier,K., Gas,S., Barry III,C.E.,  
Tekala,F., Badcock,K., Basham,D., Brown,D., Chillingworth,T.,  
Connor,R., Davies,R., Devlin,K., Feltwell,T., Gentles,S.,  
Hamlin,N., Holroyd,S., Hornsby,T., Jagsis,K., Krogh,A., McLean,J.,  
Moule,S., Murphy,L., Oliver,S., Osborne,J., Quail,M.A.,  
Rajandream,M.A., Rogers,J., Rutter,S., Seeger,K., Skellton,S.,  
Squares,S., Squires,R., Sulston,J.E., Taylor,K., Whitehead,S. and  
Barrell,E.G.

## TITLE

Deciphering the biology of Mycobacterium tuberculosis from the  
complete genome sequence

Nature 393 (6685), 537-544 (1998)  
98295987

## PUBMED

9634230

## REFERENCE

2 (bases 1 to 11700)

## AUTHORS

Parkhill,J

## TITLE

Direct Submission

## JOURNAL

Submitted (11-JUN-1998) Submitted on behalf of the Mycobacterium  
tuberculosis sequencing and mapping teams, Sanger Centre, Wellcome  
Trust Genome Campus, Hinxton, Cambridge CB10 1SA Unite de Genetique  
Moleculaire Bacterienne, Institut Pasteur, 28 rue du Docteur Roux,  
75724 Paris Cedex 15, France E-mail: parkhill@sanger.ac.uk  
On Jun 20, 1998 this sequence version replaced g1:2181960.

## COMMENT

Notes:

Details of M. tuberculosis sequencing at the Sanger Centre are  
available on the World Wide Web.  
(URL, <http://www.sanger.ac.uk/Projects/M.tuberculosis/>) CDS have  
been renumbered from the original cosmid submissions but the old  
gene designations are in brackets after the new gene numbers.  
Gene prediction was based on a Hidden Markov Model of TB genes  
implemented in TParse (Krogh) supplemented with visual inspection  
of positional base preference in codons, especially where there is  
an increase in the observed/expected third position G + C.

CAUTION: In some cases we may not have predicted the correct  
initiation codon. Where possible we choose an initiation codon  
(atg, gtg, or ttg) which is preceded by an upstream ribosome  
binding site sequence (optimally 5-13bp before the initiation  
codon). If this cannot be identified we choose the most upstream  
initiation codon.

## FEATURES

Location/Qualifiers

1..11700  
/organism="Mycobacterium tuberculosis H37Rv"  
/mol\_type="genomic DNA"  
/strain="H37Rv"  
/db\_xref="taxon:83332"  
1309..11293  
/organism="Mycobacterium tuberculosis H37Rv"  
/mol\_type="genomic DNA"  
/strain="H37Rv"  
/db\_xref="taxon:83332"  
10774..>11700  
/organism="Mycobacterium tuberculosis H37Rv"  
/mol\_type="genomic DNA"  
/strain="H37Rv"  
/db\_xref="taxon:83332"  
/clone="CI5"  
misc\_feature  
1..1409  
/note="fragment designated v031. Does not represent a  
physical clone"

83..88

## RBS

97..1674

## gene

97..1674

## CDS

97..1674

/gene="fadD7"

/note="Rv0119, (MTV031.13-MTCI418B.01), len: 525.fadD7,

PossibleCoenzyme A-ligase similar to 4-coumarate:CoA ligase of manyoraniisms e.g. PTU39405\_1 (537 aa). Contains PS00455Putative AMP-binding domain signature. FASTA scores: gpiU39405|ptu39405\_1 Pinus taeda xylem 4-coumarate:CoA (537 aa) opt:483z-score: 526.1 E(); 8.3e-22; 28.2% identity in 440 aaoverlap score is 0.896"

```

/codon_start=1
/transl_table=1
/product="fadD7"
/db_xref="GI:3242255"
/db_xref="SPTREMBL:O07169"
/translation="MASDFGRIADLVEVATRLPEAPALVVTADRIAISHRDRLARLV
DELQAGTRSLRGLPDRVALRMGNSAEFFVALLAASRADLVVPLDLPALPTEQORVS
QAAGARVYLIDADGPHDRAEPTTWMPLTVNVGDSGSPSGTSLVHLDAAVPAVPATS
TPEGLRPDAMITAGTGTGLPKVMPWTHANIASVRAIITGYRLSPDAAVPAVMPLY
HGHLASLALTAGSGTGLPARGFSAHTFWDIDKAVGATVTAAPTTHOILLERS
ATPESGRKPAALFIRSCSAPLTAQAALALQTEFAAPVCAFGMTATHQVTTQIEG
IDQETPVSTGLVGRSTGAQIRIVGSDGLPLPAGVEIWLRTGTVVRYGLGDPFIT
AANFTDGLWRTGDLGSLSAAGDLSIRGRKELINRGKEKISPERVEGVLASHPNVMEA
AVFGVPHQLYGEAAVAIVPRESAPPTREELVQFCRLAAFEIPASFOEASGLPHTA
KGSLLDRRAVAERFGHSV"
610..645
/misc_feature
/gene="fadD7"
/notes="PS00455 Putative AMP-binding domain signature,
[LIWYFY].{2}[STG][STAG]G[ST][STEI][SG].[PASLIVM][KR], info
count = 22.0"
complement(1675..3819)
/gene="fusA2"
complement(1675..3819)
/gene="fusA2"
/notes="Rv0120c, (MTCI418B.02c), fusA2, len: 714.
Elongation factor g, similar to eg. EFG_ECOLI P02996
Elongation factor g (ef-g), Escherichia coli (703 aa),
fasta scores: opt:1049 z-score: 1105.5 E(); 0..32.5%
identity in 717 aa overlap. Contains PS00017
Arp/GTP-binding site motif A (P-loop).Also similar to M.
tuberculosis fusA, MTCY210.01 (39.1% identity in 299 aa
overlap). Similar to spIP30767IEFG_MYCLE_ELONGATION FACTOR
G (EF-G). (701 aa); 31.7% identity in 710 aa overlap."
/codon_start=1
/transl_table=1
/product="fusA2"
/protein_id="CAB09448.1"
/db_xref="GI:2181962"
/db_xref="SWISS-PROT:O07170"
/translation="MADRVNASOGAAAAAPTANGPGVNRVNVLVGPSGGKTTLIEALL
VAARVLRPGSVTETGTCDFDEAIEIQRORSVGLAVASLAYDGIKVLNVDTPGVADEV
GELAGLRADCALEFVTAANEGVDEPTKSLWQECQVGMPPRAVITKLDHARANYREA
LTAQDAFGRKVLPLYLPSGDLGLISLQALYEDAGKRTRTPAESDTERIEEARGA
LIEGIIIESEDESLEMYLGGETIDESVLIQDLEKAVARGSFVPIVPCSVTGVGTLE
LLEVATRGFPSPMEHPLPEVFTPGQVPHAEACNDADAPLLAEVVVKTSDPYVGRVSLV
RVFGSTTRPDTTVHVSGHFSFEGGTSNTHPDHDEDERIGVLSFPLCKQORPAAVYV
AGDCAIGKLSRAETGDTLSPKAEPLVLPWTMPEPLLPPIAAHAKTDEDKLSVGLG
RLAADPTLRLQEQNHQVTHQVWCMEGHAHVLDLTANRYGVSDVIELRVPURETF
AGNAGHGRHILQSGHGQYGVCDIEVEPLPEGSGFEFLDKVGVGAVPRQIPNVEKG
VRAQMDKGVHAGYVPVDIRVTLTLDGKAHVSDDSFQFQAGALALREAAATKIVILLE
PIDETSVLPDPDFGVGRVIGLDSRRRGRVLGTETAGHDRTVIKAEVPPQVELTRYAIDLR
SLANGAASFTSFARYERPMPESAARVAKAGAG"
complement(3709..3732)
/misc_feature
/gene="fusA2"
/notes="PS00017 ATP/GTP-binding site motif A (P-loop)"
complement(3956..4390)
/gene="Rv0121c"
complement(3956..4390)
/gene="Rv0121c"
/notes="Rv0121c, (MTCI418B.03c), unknown, len: 144 aa"
/codon_start=1
/transl_table=1
/product="hypothetical protein Rv0121c"
/protein_id="CAB09449.1"
/db_xref="GI:2181963"
/db_xref="SPTREMBL:O07171"

```

/translation="MGEBDPKLRPAQSPVARLSTSTPDGTPHLVPPVVFALCARRPAEA  
 RDGAVYITVADKKRTTORLRRLNLEHNPRASVLVDSYADDWTQLMWVRADGVAALIH  
 TGDEVMRAAYRLLRKAYQSQSVPLNGFVIAIQVRWASWHA"  
 complement(4399..4401)  
 /note="possible RBS upstream of Rv0121c"  
 4532..4535  
 /note="possible RBS upstream of Rv0122"  
 4539..4907  
 /gene="Rv0122"  
 4539..4907  
 /gene="Rv0122"  
 /notes="Rv0122, (MTCI418B.04), unknown, len: 133 aa"  
 /codon\_start=1  
 /transl\_table=1  
 /product="hypothetical protein Rv0122"  
 /protein\_id="CAB09450.1"  
 /db\_xref="GI:2181964"  
 /db\_xref="SPTREMBL:O07172"  
 /translation="MAGVSAAAGIGWGLNVNTNRDQCYRVRTTVDALTHPEYRV  
 HTRGVQRVTRNARKHRVSKRIVAAHRHCGVPIQEDGSLYIOGHDTSGLRLTEVVA  
 VEADDGLIITHAMPKEWR"  
 4890..4895  
 /gene="Rv0122"  
 /note="possible RBS upstream of Rv0123"  
 4904..5272  
 /gene="Rv0123"  
 4904..5272  
 /gene="Rv0123"  
 /notes="Rv0123, (MTCI418B.05), unknown, len: 133 aa"  
 /codon\_start=1  
 /transl\_table=1  
 /product="hypothetical protein Rv0123"  
 /protein\_id="CAB09451.1"  
 /db\_xref="GI:2181965"  
 /db\_xref="SPTREMBL:O07173"  
 /translation="MTKKPNPADYVIGDDVEVSDVLKQEEVYVVGDERLTDERVEOM  
 ASESLRLAREANLIPGKSLSGSAHSPAVVVVSKATHAKLKLARSKMSVSKL  
 LRPVDFEYQRETRGILPRR"  
 5570..5574  
 /note="possible RBS upstream of Rv0124"  
 5581..7044  
 /gene="PE\_PGRS"  
 5581..7044  
 /gene="PE\_PGRS"  
 /note="Rv0124, (MTCI418B.06), len: 487. Member of the  
 PGRSfamily of Glycine-rich proteins, highly similar to eg.  
 YODP\_MYCTU\_Q50615 hypothetical glycine-rich 40.8 kd  
 protein (498 aa), fasta scores: opt: 1730 z-score: 1201.5  
 E(); 0, 60.7% identity in 504 aa overlap"

Alignment Scores:  
 Pred. No.: 3,52e-41 Length: 11700  
 Score: 675.00 Matches: 132  
 Percent Similarity: 100.00% Conservatative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 1 Gaps: 0

US-09-684-215A-4 (1-132) x MTCI418B (1-11700)  
 Qy 1 ThrAlaAlaSerAspAsnPhcInLeuSerGlnGlyGlnGlyPheAlaIleProfile 20  
 Db 7865 ACGCCGCGTCCGATAATTCACGCTGTCCAGGTGGCGAGGATTGCCCATTCGCGAT 7924  
 Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
 Db 7925 GGGCAGGGCATGGCGATCCGGCGCAGATCCGATCGGGTGGGGGTACCCACCGCTCAT 7984  
 Qy 41 IleGlyProThrAlaPheLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60  
 Db 7985 ATCGGGCGCTACCGCTTCTCGGCTTGGGTGCTTGTTCGACAAACACGCGACGGCGACGA 8044  
 Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80

```

|||||
8045 GTCCAAACGCTGGTGGGAGCGCTCCGGCGCAAGTCTCGCATCTCCACCGGCGGTG 8104
QY 1leThrAlaValAspGlyAlaProtleAsnSerAlaThrAlaMetAlaAsAlaLeuAsn 100
Db 8105 ATCACCAGGCTGACGGGCTCCGATCACTCGCCACCGCATGGCGGACCGCTTAAC 8164
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
Db 8165 GGGCATCATCCGCTGACGTCATCTCGGTGACCTGGCAACCAAGTCGGGCGGCACGCT 8224
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
Db 8225 ACAGGGAACGTCAGATTGGCGGAGGACCCCGCC 8260

RESULT 5
LOCUS AE006925 14029 bp DNA linear BCT 27-APR-2001
DEFINITION Mycobacterium tuberculosis CDC1551, section II of 280 of the
complete genome.
ACCESSION AE006925 AE000516
VERSION AE006925.1 GI:13879610
KEYWORDS
SOURCE Mycobacterium tuberculosis CDC1551
ORGANISM Mycobacterium tuberculosis CDC1551
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
REFERENCE 1 (bases 1 to 14029)
AUTHORS Fleischmann,R.D., Alland,D., Eisen,J.A., Carpenter,L., White,O.,
Peterson,J., DeBoy,R., Dodson,R., Gwinn,M., Haft,D., Hickey,E.,
Kolony,J.F., Nelson,W.C., Umayam,L.A., Ermolaeva,M.,
Salzberg,S.L., Delcher,A., Utterback,T., Weidman,J., Khouri,H.,
Gill,J., Mikula,A. and Bishai,W.
TITLE Whole genome comparison of Mycobacterium tuberculosis clinical and
laboratory strains
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 14029)
AUTHORS Fleischmann,R.D., Alland,D., Eisen,J.A., Carpenter,L., White,O.,
Peterson,J., DeBoy,R., Dodson,R., Gwinn,M., Haft,D., Hickey,E.,
Kolony,J.F., Nelson,W.C., Umayam,L.A., Ermolaeva,M.,
Salzberg,S.L., Delcher,A., Utterback,T., Weidman,J., Khouri,H.,
Gill,J., Mikula,A. and Bishai,W.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2001) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
FEATURES
Source
1..14029
Location/Qualifiers
/organism="Mycobacterium tuberculosis CDC1551"
/mol_type="genomic DNA"
/strain="CDC1551"
/db_xref="taxon:83331"
/note="clinical strain"
complement(99..533)
/gene="MT0129"
complement(99..533)
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK44351.1"
/db_xref="GI:13879611"
/transl_table="MGFFDKLRFQAQSPVARLATSTPDGTPHLVPVVFALGARRPAEA
TGADVIYAVDAKRTTQRLRLANLEHNPASVLDYSADDTQLWVRADGVAAIH
RDGVMRAATLLRAKYTOQSVPLNGPVIAIVQRASRAH"
complement(643..1062)
/gene="MT0130"
complement(643..1062)
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="alpha-amylase family protein"
/protein_id="AAK44358.1"
/db_xref="GI:13879616"
/transl_table="MNEAHSVEHPVQGSHEVGVVEHPDAKDFGSAALPADPTWF
KHAFYEVLYRAFFDASADSGDLRLIDRLDYQLWLGLDCLWLPFPVDSPLRGGVD
IRDFYKVLPEFGTVDDDFVALDAHRGIRITLDVMNHTSESHWQESRRDPGY
GDYVWSDTSERYTDARIIFVDTEESNWSDEPVRQRQFTHRRFFSHQDPLNDPAVQE
AMIDVIRFWLGLDGRFLDAVPYLFEREECTNCLPETHAFKLRKRVVDDEFPGRV

/product="hypothetical protein"
/protein_id="AAK44354.1"
/db_xref="GI:13879612"
/transl_table="MASSSLPLLRHCVSDDVTVVGVGDGDDLGKCTARRIAALVVQRA
IFLNDRTVAHSCDDAYLGHAVLPVGPDPDPDLHASSMYSLCMQSVNCRPPDATA
LVSVRLCHVOTDPTDSCGRDRPGQLPCAPLDYHRRH"
1047..1415
/gene="MT0131"
1047..1415
/gene="MT0131"
1047..1415
/note="identified by match to protein family HMM"
/codon_start=1
/transl_table=11
/product="DNA-binding protein, CopG family"
/protein_id="AAK44355.1"
/db_xref="GI:13879613"
/transl_table="MTKPRNPADYVIGDVEVSDVLKOEVVVDGERLTDERVEOM
ASESLRLAREANLIPGKSLSGSAHSPAVQVVVSKATHAKELARSKMSVSKL
LRPVLEDFVQRETRILPRR"
1682..3367
/gene="MT0132"
1682..3367
/gene="MT0132"
1682..3367
/note="similar to GB:AL123456; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="PE_PGRS family protein"
/protein_id="AAK44356.1"
/db_xref="GI:13879614"
/transl_table="MFAGGAGGGLRCVMSFVSAPEIVAAATDLAGIGSAISAANA
AAAATPVAAGADEVSAATAALFSGHAQYALSAQAFAHQFVQTLTAGGAGAYA
AAEAQVEQQLIAAINAPTALLRLPGINGADGAPGTQAGAGGILYGNNGSGSA
AGGAQAGGAPAGLIGHGSGGAGGAGGAGGHWLWNGVGGGAGGAGVAG
GHGAGCAAGLWAGGAGGAGGAGGADANIVSGDGLGGAGGGMWLYDGGAGGHG
QGAGGAGGAGGCGGAGRGLWGTGCGAGHGCGGTGPPPLPGAGMGAAGAGGAG
LIGNGAGGAGGAGGAGGAGVAGVXGAGGNAMLIHGAGGAGGAGGAGGAGGAGGAG
PLFGNGGAGGAGGAGVAGTGNITGGAGGAGGAGDARLIGHGAGGAGGAGGAGGAG
PGHNGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
3519..4586
/gene="MT0133"
3519..4586
/gene="MT0133"
/note="identified by match to protein family HMM"
/codon_start=1
/transl_table=11
/product="serine protease, putative"
/protein_id="AAK44357.1"
/db_xref="GI:13879615"
/transl_table="MSNSRRSLRWSMLLSVLAAVGLGLATAPAAQAPALSDORFAD
FPALPLDPSAMVAQVGPVNVNITKLGYNNAVAGTGVIPDNGVWLTHNVIAGATD
INAFVSGGTYGVYVGYDPTDVAVLQLRGAGLPSAATGGGVAVGEPVAVNGNSG
QGCTPRAPGRVVALQCTVOASDLSLGAETLNLGLOFDDAAIOFGSGGPGVNLGQ
VWGMTAASDNFQSLGQGGQFAIPIGQAMAIAQIRSGGSGPTVHIGTAFGLGVVD
NNGNARVQRVVGSAAPASLGIISTGDIITAVDGPAPINSATAMADALNHHHPGDIISVT
WQTKSGGTRTGNVTLAEGPPA"
4695..6500
/gene="MT0134"
4695..6500
/gene="MT0134"
/note="similar to GP:2808807; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="alpha-amylase family protein"
/protein_id="AAK44358.1"
/db_xref="GI:13879616"
/transl_table="MNEAHSVEHPVQGSHEVGVVEHPDAKDFGSAALPADPTWF
KHAFYEVLYRAFFDASADSGDLRLIDRLDYQLWLGLDCLWLPFPVDSPLRGGVD
IRDFYKVLPEFGTVDDDFVALDAHRGIRITLDVMNHTSESHWQESRRDPGY
GDYVWSDTSERYTDARIIFVDTEESNWSDEPVRQRQFTHRRFFSHQDPLNDPAVQE
AMIDVIRFWLGLDGRFLDAVPYLFEREECTNCLPETHAFKLRKRVVDDEFPGRV

```

gene  
CDS  
L1AENQPGDVEYFGDNTGGDECHMAFHPLMPRIAMVAVRRESRPISIIAOTP  
PIPDMAQMGIFURNHDELFTLENTDEEDYMYAEYAKOPRMKANTGIRRRAPLILND  
RQIIEFTALLISLSPVLYYGDEIGMDVILGDRGVIRPOMWTDPNAGFSTAN  
PGRLLYPSPQDVPVGOAVNVEAORDTSTSLNFTRLAVRRRHPAFAVGAFOELGG  
SNPSVLAVVROVAGDDGDTVLVCNNLSRPPQFIELDLQONTNTPVELTGHVEFRIG  
QVYLLTLPCHGFYWFQLTTHVEGAPPTCGGERLL  
6603..7970  
/gene="MT0135"  
6603..7970  
/gene="MT0135"  
/note="similar to GP:2808801; identified by sequence  
similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="pep2 protein"  
/protein\_id="AAK44359.1"  
/db\_xref="GI:13879617"  
/translation="MTRSDATLTKLPWSDWLPQRWYAGNRRELATVTPGVVVALRHN  
LDLVLVDVTGATERYQVLGWDFEASEYGTAAIGVADDRGTGFDAIYDVAGPQF  
LISLVSSAVCTSGTEVTREPDVELFPAAQPRVCDAEQSNFTGFDRRAILKVF  
RVSSGIPNIDELNRLVTRAGNPHVLLGAYOFGPRNRSPTDALAYALGMVTEYANA  
AGWAMATSVRDLFAEGDLYAHEVGGDFAGESYRLGEAVASVHATLADSLGTAATF  
PVDMLARLSSVAVVPELREYAPTEIQQFQKLAEAITVQVHGDHLGLQVLRTPES  
WLLIDFEGEGPDLERRAPDSPLADVAGVLRSEFYAATGPDQATDKQLAARAW  
VERNRAACDGYAVASGIDPRDSALLLGGAYELDKAVYETGYETRRPGLPILRSIA  
RITAS"  
8038..8817  
/gene="MT0136"  
8038..8817  
/gene="MT0136"  
/note="identified by Glimmer2; putative"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="AAK44360.1"  
/db\_xref="GI:13879618"  
/translation="MQREIYDEARLSVLAALAGILGATATFTHSAGVFTVMTGNSQ  
RAVLGFGDAMWSTFALLIFFVAGVVIASVCRHFHWAHPHGTVLTTFSLIFAA  
GYDILMGHSMLEDFPILFVFGIGALNTSFVKDGEVSPVSTGTVLKMGOGIE  
RHLAGKGVDMWGLFLLHASFVLGAAGAGISMTVTGPOMLAVAAVVCAATTGYTYLH  
ADRGLVGNKRPQPKRLFRALRRGELDSGTSTPATNYGSS"  
complement(8949..9989)  
/gene="MT0137"  
complement(8949..9989)  
/gene="MT0137"  
/note="similar to SP:P31953 GB:X57229 PID:48828  
PID:1567734 PID:1877254; identified by sequence  
similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="esterase, putative, antigen 85-c"  
/protein\_id="AAK44361.1"  
/db\_xref="GI:13879619"  
/translation="MGLVVMVTFEQVRLRSLAATLPRLLAIAAMGAVLVVGLVGF  
GGPATAFASREPLVEYLQVPSASMDRIKQVFOGGPGRAYLLDGLRAQDDYNGWD  
INTPAFEYQSGLSVIMPVGGQSFYDQVPSOSNGQNTYKWEITREMPALQ  
ANKGVSPNAAVGLSMGSGSALIAAYPOQFPYAAASLSGLNSEGWPTLLGLAM  
NDSGGYNANMGPSSDPAWKRNDPMQVIRLVANTRIIWYCGNGTSPDLGDNIPA  
KELEGTLRTNTFRDTVAADGGRNGVFNPPNGTSHSPYNNQVLVANKADIQHVING  
ATPPAAPAAPAA"  
10218..10673  
/gene="MT0138"  
10218..10673  
/gene="MT0138"  
/note="similar to SP:P08634 PID:46222; identified by  
sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="modulation protein N-related protein"  
/protein\_id="AAK44362.1"  
/db\_xref="GI:13879620"

Alignment Scores:

Pred. No.: 4,21e-41 Length: 14029  
Score: 675.00 Matches: 132  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 0 Gaps: 0  
US-09-684-215A-4 (1-132) x AE006925 (1-14029)  
Qy 1 ThrAlaAlaSerAspAspPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20  
Db 4188 ACGGCGCGGTGCGATAACTTCCAGCTGTCCAGGTGGGCGAGGATCCCGCATTCGGATC 4247  
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
Db 4248 GGGCAGGCGATGGCGATCGCGGCCAGATCCGATCGGTGGGGGTTCACCCACCGTTCAT 4307  
Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArg 60  
Db 4308 ATCGGCGCTACGGCTTCCTCGGCTGGGTGTTCGACACACACGCGACGCGACGA 4367  
Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 4368 GTCCACCGCGGTGGTGGGAGCGCTCCGCGCGAAGTCTCGGCACTCTCCACCGCGACGTG 4427  
Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAlaLeuAsn 100  
Db 4428 ATCACCAGCGGTGGCGCGGCTCCGATCACTCGGCCACCGCATGGCGACGCGCTTAAC 4487  
Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
Db 4488 GGGCATCATCCGGGTGACGTCTCTCGGTGACCTGGCAACCAAGTCGGCGCGACGCGT 4547  
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
Db 4548 ACAGGAGCGTGCATATGGCCGAGGACCCCGGCC 4583  
RESULT 6  
BX248334  
LOCUS  
DEFINITION  
Mycobacterium bovis subsp. bovis AF2122/97 complete genome; segment 1/14  
ACCESSION  
BX248334 BX248333  
VERSION  
BX248334.1 GI:31616762  
KEYWORDS  
complete genome.  
SOURCE  
Mycobacterium bovis subsp. bovis AF2122/97  
ORGANISM  
Mycobacterium bovis subsp. bovis AF2122/97  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium  
tuberculosis complex.  
REFERENCE  
1  
AUTHORS  
Garnier, T., Elglmeier, K., Camus, J.-C., Medina, N., Mansoor, H.,  
Przyor, M., Duthoy, S., Grondin, S., Lacroix, C., Monsemp, C., Simon, S.,  
Harris, B., Atkin, R., Doggett, J., Mayes, R., Keating, L.,  
Wheeler, P., Parkhill, J., Barrell, B.G., Cole, S.T., Gordon, S.V. and  
Hewinson, G.  
TITLE  
The complete genome sequence of Mycobacterium bovis  
JOURNAL  
Online Publication  
REMARK  
PNAS 10.1073/pnas.1130426100 ( Microbiology )  
REFERENCE  
2 (bases 1 to 1343050)  
AUTHORS  
Garnier, T.  
DIRECT SUBMISSION  
TITLE  
Submitted (24-MAR-2003) Garnier T., Unite de Genetique Moleculaire  
Bacterienne Institut Pasteur 28, rue du Dr Roux 75724 PARIS cedex  
15, France. e-mail:tgarnier@pasteur.fr Submitted on behalf of the  
Mycobacterium bovis sequencing teams, TB Research Group, Veterinary  
Laboratories Agency Weybridge, Woodham Lane, New Haw, Addlestone,  
Surrey KT15 3NB, UK. Sanger Centre, Wellcome Trust Genome Campus,  
Hinxton, Cambridge CB10 1SA, UK. Pf4 Annotation, Genopole, Institut  
Pasteur, 28 Rue du Docteur Roux, 75724 Paris Cedex 15, France.  
Unite de Genetique Moleculaire Bacterienne, Institut Pasteur, 28  
rue du Docteur Roux, 75724 Paris Cedex 15, France  
LOCATION/Qualifiers  
FEATURES

```

source      1. .343050
/organism="Mycobacterium bovis subsp. bovis AF2122/97"
/mol_type="genomic DNA"
/strain="AF2122/97"
/db_xref="taxon:233413"
1. .1524
/gene="dnaA"
/locus_tag="Mb0001"
CDS        1. .1524
/gene="dnaA"
/locus_tag="Mb0001"
/locus_tag="Mb0001"
/locus_tag="Mb0001"
len: 507 aa. Equivalent to Rv0001,
(99.6% identity in 507 aa overlap). dnaA, chromosomal
replication initiator protein (see citations below),
equivalent to other Mycobacterial CHROMOSOMAL REPLICATION
INITIATOR PROTEINS e.g. P46388|DNAA_MYCLE from
Mycobacterium leprae (502 aa); Q9L7L7|DNAA_MYCPA from
Mycobacterium paratuberculosis (509 aa); P49990|DNAA_MYCAV
from Mycobacterium avium (508 aa); P49992|DNAA_MYCSM from
Mycobacterium smegmatis (504 aa); etc. Also highly similar
to others except in N-terminus e.g. Q9ZH75|DNAA_STRCH
CHROMOSOMAL REPLICATION INITIATOR PROTEIN from
Streptomyces chrysomallus (624 aa); Q9ZH76|DNAA_STRRE from
Streptomyces reticuli (643 aa); DNAA_ECOLI|P03004|B3702
chromosomal replication initiator protein from Escherichia
coli strain K12 (467 aa). FASTA scores: opt: 986, E(): 0,
(43.2% identity in 389 aa overlap); etc. Contains P500017
ATP/GTP-binding site motif A (P-loop) and P501008 DnaA
protein signature. BELONGS TO THE DNAA FAMILY. Note that
the first base of this gene has been taken as base 1 of
the Mycobacterium bovis genomic sequence.
/codon_start=1
/evidence=experimental
/transl_table=11
/product="CHROMOSOMAL REPLICATION INITIATOR PROTEIN DNAA"
/protein_id="CAD92863.1"
/db_xref="GI:31616763"
/translation="MTDDPGSGFTTVMNAVVELNGDKPKVDGPPSSDANLSAPLTPQO
RAWNLVQPLTIVGTFALLSVFQNEIEHRLRAPITDALSRLGHGHOIQGVRIAP
PATRDADDTVPSPENPATTSDTTDNDDEIDDSAAAGDNOHSPVETERPNTDS
ATAGVTSLNRRYTFEDTFVIGASNRFAHAAALAEAPARAYNPLFWGESGLGKTHLL
HAAGNVAQKLFPGMRVYVSTEEFNDFNSRDOKKFAKRSYDQVLLVDDIQFI
EKEKGIEQFFHTTHNANKOIVISSDRPKPKLATLEDRLTRFEWGLINDTVQPE
LETRAILKKAQMERLAIPDDVLELIASSIERNTRELEGALRTVFAFASLNTIDIK
ALAEIVRLDIADANTMQISAATIMAAATAEYDTTVEELRGPGKTRALAAQSQIAMYL
CRELTDLSLPKIQAGFRDHTTVMVAQRKILSEMAERREVDHVKELTRIQRSKR"
2052. .3260
/gene="dnaN"
/locus_tag="Mb0002"
CDS        2052. .3260
/gene="dnaN"
/locus_tag="Mb0002"
/locus_tag="Mb0002"
/locus_tag="Mb0002"
len: 402 aa. Equivalent to Rv0002,
(99.8% identity in 402 aa overlap). dnaN, DNA polymerase
III (beta chain) (EC 2.7.7.7) (see citations below),
equivalent to other Mycobacterial DNA POLYMERASES III BETA
CHAIN e.g. NP_301130.1|NC_002677 from Mycobacterium leprae
(399 aa); Q9L7L6|DP3B_MYCPA from Mycobacterium avium
subsp. paratuberculosis (399 aa); P52851|DP3B_MYCSM from
Mycobacterium smegmatis (397 aa); etc. Also highly similar
to others e.g. P27903|DP3B_STRCO DNA POLYMERASE III BETA
CHAIN from Streptomyces coelicolor (376 aa). FASTA scores:
opt: 1189, E(): 0, (52.8% identity in 337 aa overlap);
P21174|DP3B_MICUJ from Micrococcus luteus (310 aa);
P52073|DP3B_SYNPF from Synecococcus sp. strain PCC 7942
(375 aa); etc. Overlaps and extends CDS in neighbouring
cosmid MFCY10H4.01."
/codon_start=1
/evidence=experimental
/transl_table=11
/product="DNA POLYMERASE III (BETA CHAIN) DNAA (DNA
NUCLEOTIDYLTRANSFERASE)"
/protein_id="CAD92864.1"
/db_xref="GI:31616764"
/translation="MDAATRVGLTDLTFLRLRESFADAVSVAKNLPARPAPVPLSG
VLTGSDNGLTISGDEYVSAQAQVAGVSGSVLYSGRLSDITRALPNKPVGVHV
EGNRVALTGCGNARFSLPTMPVEDYPTPLTBETGLLPALFAELVPAKTLAAGRDDT
LPMLTGIRVEILGETVLAATDRFLAVRELKWSASSPDIEAALVPAKTLAAGRAKAG
IGGSDVRLSGTGPVGKDLGSGNGKRSSTRLLIDAEFFKROLPLPHTFAVATMD
VABLTAIKLVALVADRGAGVMEFADGVRVLSAGADDDVRAEEDLVVDYAGEPLTIA
FNPTVLTGCLSSLSRSERVSFGFTTAGKPAALLRPVSGDDRPVAGLNGGFPFAVSTDYV
YLLMPVRLPG"
3280. .4437
/gene="recF"
/locus_tag="Mb0003"
CDS        3280. .4437
/gene="recF"
/locus_tag="Mb0003"
/locus_tag="Mb0003"
/locus_tag="Mb0003"
len: 385 aa. Equivalent to Rv0003,
(99.5% identity in 385 aa overlap). recF, DNA replication
and repair protein (see citations below), equivalent to
others Mycobacterial DNA replication and repair proteins
e.g. NP_301131.1|NC_002677 from Mycobacterium leprae (385
aa); Q9L7L5|RECF_MYCPA from Mycobacterium avium subsp.
paratuberculosis (385 aa); P50916|RECF_MYCSM from
Mycobacterium smegmatis (384 aa); etc. Also highly similar
to others e.g. P36176|RECF_STRCO DNA REPLICATION AND
REPAIR PROTEIN from Streptomyces coelicolor (373 aa);
NP_440892.1|NC_000911 from Synecocystis sp. strain PCC
5803 (384 aa); NP_469352.1|NC_003212 from Listeria innocua
(370 aa); etc. Contains P50017 ATP/GTP-binding site motif
A (P-loop), P500617 RecF protein signature 1, and P500618
RecF protein signature 2. BELONGS TO THE REC F FAMILY."
/codon_start=1
/evidence=experimental
/transl_table=11
/product="DNA REPLICATION AND REPAIR PROTEIN REC F
(SINGLE-STRAND DNA BINDING PROTEIN)"
/protein_id="CAD92865.1"
/db_xref="GI:31616765"
/translation="MYVRHLGLDRFSWACVDLELHPGRTVFVPGNGYKTNLIELM
YSTTLGSHVSADLPLIRVGTDRAVISTIVNDGRECAVDLEIATGRVKNRKNRSSV
RSTDGVVLRVLAFLPFLRVDGPDADRRYLDLAIIVRRPATVAARVYERVYRQ
RTALKSVPGRVGRDGVFDLEVMSRLAEHGAELVAARDLQNLAEVKKAYOL
LAPESRSASIGYRSMVDVTGPSEQSDTROLLAARLLAALARRDLERGVCLVGP
RDLLILRGDQPAKGFASHGEASLVALRLAAYOLLRVGDGEPVLLDDVPAELDMV
RRALATAESAEQVLTVAALVEDIPAGHDARVHIDVRADDTGSGSVSLP"
4434. .4997
/locus_tag="Mb0004"
CDS        4434. .4997
/locus_tag="Mb0004"
/locus_tag="Mb0004"
/locus_tag="Mb0004"
len: 187 aa. Equivalent to Rv0004, len:
187 aa, from Mycobacterium tuberculosis strain H37Rv,
(99.5% identity in 187 aa overlap). Conserved hypothetical
protein (see citation below), highly similar, but longer
21 aa in N-terminus, to AAF33696.1|AF222789 unknown
protein from Mycobacterium avium subsp. paratuberculosis
(166 aa); and highly similar to NP_301132.1|NC_002677
conserved hypothetical protein from Mycobacterium leprae
(189 aa); S70990 hypothetical protein from Mycobacterium
smegmatis (194 aa). Also similar to in C-terminus to
C-terminal part of P35925|YREG_STRCO HYPOTHETICAL 19.8 KDA
PROTEIN (IN REC F-GYRB INTERGENIC REGION) from Streptomyces
coelicolor (190 aa). FASTA scores: opt: 404, E(): 3.9e-18,
(40.7% identity in 189 aa overlap).
/codon_start=1
/transl_table=11
/product="CONSERVED HYPOTHETICAL PROTEIN"
/protein_id="CAD92866.1"
/db_xref="GI:31616766"
/translation="WTGSDVDRPDQNGERLMSKSPGLDLVRRPTLDEARAAARGQDAG
RGRVASVSGRVAGRRRSWSGPGPDIRDPPQLGKAARELAKRGRSVRVAEGVILGOW

```

SAVVGHOIAEHARTALNDGVLVSIAESTAWATOLRIMQALAKIAAAVNDVVRSL  
KITGPAPASWRKGRPHIAGRPRDTYG"

gene

/gene="gyrB"  
/locus\_tag="Mb0005"  
5123..7267

CDS

/gene="gyrB"  
/locus\_tag="Mb0005"  
/EC\_number="5.99.1.3"

len: 714 aa. Equivalent to Rv0005.  
/note="Mb0005, gyrB, len: 714 aa. Equivalent to Rv0005.  
(99.9% identity in 714 aa overlap). gyrB, DNA gyrase  
subunit B (EC 5.99.1.3) (see citations below), equivalent,  
except in N-terminus, to other Mycobacterial DNA GYRASES  
SUBUNIT B e.g. T10005 from Mycobacterium leprae (697 aa);  
Q9L7L3(GYRB\_MYCPA from Mycobacterium avium subsp.  
paratuberculosis (677 aa) (has its N-terminus shorter);  
P48355(GYRB\_MYCSM from Mycobacterium smegmatis (675 aa);  
etc. Also highly similar to others e.g. T10969 from  
Streptomyces coelicolor (686 aa); P50075(GYBS\_STRSH from

## Alignment Scores:

Pred. No.: 1.01e-39 Length: 343050  
Score: 675.00 Matches: 132  
Percent Similarity: 100.00% Conservativeness: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 1 Gaps: 0

US-09-684-215A-4 (1-132) x BX248334 (1-343050)

Qy 1 ThrAlaLaSerAspAsnPhleGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20

Db 152008 ACGGCGGTCGCGATACTTCCAGCTGTCCAGGTCGGGAGGATTCGCCATTCGATC 152067

Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnLeuSerGlyGlySerProThrValHis 40

Db 152068 GGGCAGGCGATGGCGATCGCGGCGCAGATCCGATCGGGTGGGGTCCACCACCGTTTCAT 152127

Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60

Db 152128 ATCGGGCTACCGCTTCTCGGCTTGGTGTTCGACACAAACGCAAGCGGCGACGTG 152187

Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80

Db 152188 GTCCAACGCGTGGTCCGAGCGCTCCGCGGCAAGTCTCCGCATCTCCACCGGCGACGTG 152247

Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100

Db 152248 ATCACCAGGTCGACGGCGTCCGATCACTCGGCCACCGCATGGCGGCGCGCTTAAC 152307

Qy 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120

Db 152308 GGGCATCATCCGCTGACGTATCTCGTGACCTGGCAACCAAGTCCGCGGCGCACGGT 152367

Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132

Db 152368 ACAGGGAACGTGACATTTGGCGGAGGACCCCGGCC 152403

RESULT 7

LOCUS AR169152

DEFINITION Sequence 4 from patent US 6290969.

ACCESSION AR169152

VERSION AR169152.1 GI:17906927

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 447)

AUTHORS Reed, S.G., Skeiky, Y.A.W., Dillon, D.C., Campos-Neto, A., Houghton, R.,

TITLE Compounds and methods for immunotherapy and diagnosis of

tuberculosis

Patent: US 6290969-A 4 18-SEP-2001;

FEATURES

Location/Qualifiers  
1..447

BASE COUNT

79 a 146 c 149 g 72 t 1 others

ORIGIN

Alignment Scores:

Pred. No.: 3.24e-42 Length: 447  
Score: 670.00 Matches: 131  
Percent Similarity: 99.24% Conservativeness: 0  
Best Local Similarity: 99.24% Mismatches: 1  
Query Match: 99.26% Indels: 0  
DB: 6 Gaps: 0

US-09-684-215A-4 (1-132) x AR169152 (1-447)

Qy 1 ThrAlaLaSerAspAsnPhleGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20

Db 11 ACGGCGGTCGCGATACTTCCAGCTGTCCAGGTCGGGAGGATTCGCCATTCGATC 70

Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnLeuSerGlyGlySerProThrValHis 40

Db 71 GGGCAGGCGATGGCGATCGCGGCGCAGATCCGATCGGGTGGGGTCCACCACCGTTTCAT 130

Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60

Db 131 ATCGGGCTACCGCTTCTCGGCTTGGTGTTCGACACAAACGCAAGCGGCGACGTG 190

Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80

Db 191 GTCCAACGCGTGGTCCGAGCGCTCCGCGGCAAGTCTCCGCATCTCCACCGGCGACGTG 250

Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100

Db 251 ATCACCAGGTCGACGGCGTCCGATCACTCGGCCACCGCATGGCGGCGCGCTTAAC 310

Qy 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120

Db 311 GGGCATCATCCGCTGACGTATCTCGTGACCTGGCAACCAAGTCCGCGGCGCACGGT 370

Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132

Db 371 ACAGGGAACGTGACATTTGGCGGAGGACCCCGGCC 406

## RESULT 8

LOCUS AR182442

DEFINITION Sequence 4 from patent US 6338852.

ACCESSION AR182442

VERSION AR182442.1 GI:20225649

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 447)

AUTHORS Reed, S.G., Skeiky, Y.A.W., Dillon, D.C., Campos-Neto, A., Houghton, R.,

TITLE Compounds and methods for diagnosis of tuberculosis

JOURNAL

FEATURES

Location/Qualifiers  
1..447

BASE COUNT

79 a 146 c 149 g 72 t 1 others

ORIGIN

Alignment Scores:

Pred. No.: 3.24e-42 Length: 447  
Score: 670.00 Matches: 131  
Percent Similarity: 99.24% Conservativeness: 0  
Best Local Similarity: 99.24% Mismatches: 1  
Query Match: 99.26% Indels: 0

```
DB: 6 Gaps: 0
US-09-684-215A-4 (1-132) x AR182442 (1-447)
Qy 1 ThrAlaIaSerAspAsnPhelGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20
Db 11 ACGCCGGCTCCGATACCTCCAGCTGTCCAGGGTGGCAGGATTCGCCATTCCTCCGATC 70
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
Db 71 GGGCAGCGATGGCGATCGCGGGCCAGATCCGATCCGATCGGGGTACACCGCTTCAT 130
Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAlaArg 60
Db 131 ATCGGGCTACCGCTTCTCGGCTTGGTGTTCGACAAACACGGCAACGGCGCACGA 190
Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
Db 191 GTCCAACCGTGGTGGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGCGACGTG 250
Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
Db 251 ATCACCAGCGTCCGATACCTCCAGCTCCGATCACTCGGTGAACCTGGCAACAACTCGGGCGCACGGT 310
Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyClyThrArg 120
Db 311 GGGCATCATCCGCTTCTCGGCTTGGTGTTCGACAAACACGGCAACGGCGCACGGT 370
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
Db 371 ACAGGAACGTGACATTTGGCCGAGGACCCCGGCC 406
RESULT 9
AR194825
LOCUS AR194825 447 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 4 from patent US 6350456.
ACCESSION AR194825
VERSION AR194825.1 GI:20244262
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 447)
AUTHORS Reed,S.G., Skeiky,Y.A.W. and Dillon,D.C.
TITLE Compositions and methods for the prevention and treatment of M.
JOURNAL tuberculosis infection
FEATURES
source
Location/Qualifiers
1..447
/organism="unknown"
BASE COUNT 79 a 146 c 149 g 72 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 3.24e-42 Length: 447
Score: 670.00 Matches: 131
Percent Similarity: 99.24% Conservative: 0
Best Local Similarity: 99.24% Mismatches: 1
Query Match: 99.26% Indels: 0
DB: 6 Gaps: 0
US-09-684-215A-4 (1-132) x AR194825 (1-447)
Qy 1 ThrAlaIaSerAspAsnPhelGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20
Db 11 ACGCCGGCTCCGATACCTCCAGCTGTCCAGGGTGGCAGGATTCGCCATTCCTCCGATC 70
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
Db 71 GGGCAGCGATGGCGATCGCGGGCCAGATCCGATCCGATCGGGGTACACCGCTTCAT 130
Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAlaArg 60
Db 131 ATCGGGCTACCGCTTCTCGGCTTGGTGTTCGACAAACACGGCAACGGCGCACGA 190
Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
Db 191 GTCCAACCGTGGTGGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGCGACGTG 250
Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
Db 251 ATCACCAGCGTCCGATACCTCCAGCTCCGATCACTCGGTGAACCTGGCAACAACTCGGGCGCACGGT 310
Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyClyThrArg 120
Db 311 GGGCATCATCCGCTTCTCGGCTTGGTGTTCGACAAACACGGCAACGGCGCACGGT 370
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
Db 371 ACAGGAACGTGACATTTGGCCGAGGACCCCGGCC 406
RESULT 9
AR194825
LOCUS AR194825 447 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 4 from patent US 6350456.
ACCESSION AR194825
VERSION AR194825.1 GI:20244262
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 447)
AUTHORS Reed,S.G., Skeiky,Y.A.W. and Dillon,D.C.
TITLE Compositions and methods for the prevention and treatment of M.
JOURNAL tuberculosis infection
FEATURES
source
Location/Qualifiers
1..447
/organism="unknown"
BASE COUNT 79 a 146 c 149 g 72 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 3.24e-42 Length: 447
Score: 670.00 Matches: 131
Percent Similarity: 99.24% Conservative: 0
Best Local Similarity: 99.24% Mismatches: 1
Query Match: 99.26% Indels: 0
DB: 6 Gaps: 0
US-09-684-215A-4 (1-132) x AR233097 (1-447)
Qy 1 ThrAlaIaSerAspAsnPhelGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20
Db 11 ACGCCGGCTCCGATACCTCCAGCTGTCCAGGGTGGCAGGATTCGCCATTCCTCCGATC 70
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
Db 71 GGGCAGCGATGGCGATCGCGGGCCAGATCCGATCCGATCGGGGTACACCGCTTCAT 130
Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAlaArg 60
Db 131 ATCGGGCTACCGCTTCTCGGCTTGGTGTTCGACAAACACGGCAACGGCGCACGA 190
Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
Db 191 GTCCAACCGTGGTGGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGCGACGTG 250
Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
Db 251 ATCACCAGCGTCCGATACCTCCAGCTCCGATCACTCGGTGAACCTGGCAACAACTCGGGCGCACGGT 310
Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyClyThrArg 120
Db 311 GGGCATCATCCGCTTCTCGGCTTGGTGTTCGACAAACACGGCAACGGCGCACGGT 370
```



```

QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
Db 371 ACAGGGAAGTGCATTTGGCCGAGGACCCCGGCC 406

RESULT 11
AX429596
LOCUS AX429596 447 bp DNA linear PAT 21-JUN-2002
DEFINITION Sequence 4 from Patent EP1203817.
ACCESSION AX429596
VERSION AX429596.1 GI:21540845
KEYWORDS
SOURCE unidentified
ORGANISM unclassified.

REFERENCE
1 Reed, S.G., Skeiky, Y.A., Dillon, D.C., Campos-Neto, A., Houghton, R.L.,
AUTHORS Vedvick, T.S., and Twardzik, D.R.
TITLE Compounds and methods for immunotherapy and diagnosis of
JOURNAL tuberculosis
CORIXA CORPORATION (US)
FEATURES
source
1..447
Location/Qualifiers
BASE COUNT 79 a 146 c 149 g 72 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 3,248-42 Length: 447
Score: 670.00 Matches: 131
Percent Similarity: 99.24% Conservative: 0
Best Local Similarity: 99.24% Mismatches: 1
Query Match: 99.26% Indels: 0
DB: 6 Gaps: 0

US-09-684-215A-4 (1-132) x AX429596 (1-447)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 11 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGTGGCGAGGATTCGCCATTCGGATC 70

QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40
Db 71 GGGCAGGGGATGGCGATCGCGGCCAGATCCGATCGGTGGGGGTCAACCCACCGTTTCAT 130

QY 41 IleGlyProThrAlaPheLeuGlyValValAspAsnGlyAsnGlyAlaArg 60
Db 131 ATCGGCGCTTACCGCTTCTCGGCTGGGTGTTCGACAAACACGCGGCGCAGCA 190

QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
Db 191 GTCCAAACGGTGGTGGGAGCGCTCCGCGGCAAGTCTCGGATCTCCACCGCGCAGTG 250

QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
Db 251 ATCAGCGGGTGGAGCGGCTCCGATCACTCGGCAACCGCGATCGCGGCGCAGCGT 310

QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyThrArg 120
Db 311 GGGCATCATCCCGGTGACGTCTATCTCGTGAACCTGGCAACCAACGATCGGGCGGCGCGGT 370

QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
Db 371 ACAGGGAAGTGCATTTGGCCGAGGACCCCGGCC 406

RESULT 12
BD006325
LOCUS BD006325 447 bp DNA linear PAT 31-JAN-2002
DEFINITION Compounds and methods for diagnosis of Tuberculosis.
ACCESSION BD006325
VERSION BD006325.1 GI:18634696

```

```

KEYWORDS JP 2001500383-A/4.
SOURCE unidentified
ORGANISM unclassified
REFERENCE 1 (bases 1 to 447)
AUTHORS Reed, S.G., Skeiky, Y.A.W., Dillon, D.C., Neto, A.C., Houghton, R.,
TITLE Vedvick, T.S., Twardzik, D.R. and Lodes, M.J.
JOURNAL Compounds and methods for diagnosis of Tuberculosis
COMMENT Patent: JP 2001500383-A 4 16-JAN-2001;
CORIXA CORP
OS Unidentified
PN JP 2001500383-A/4
PD 16-JAN-2001
PF 07-OCT-1997 JP 1998518432
PR 11-OCT-1996 US 08/729622.13-MAR-1997 US 08/818111 PI
STEVEN G REED, YASIR A W SKEIKY, DAVIN C DILLON, PI ANTONIO CAMPOS
NETO, RAYMOND HOUGHTON, THOMAS S VEDVICK, DANIEL R THWARDZIK, PI
MICHAEL J LODES
PC C12N15/31, C07K14/35, C07K16/12, C12Q1/68, C12N15/62, G01N33/53 CC
Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT source 1..447
/organism='Unidentified'.
FEATURES
source
1..447
Location/Qualifiers
BASE COUNT 79 a 146 c 149 g 72 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 3,248-42 Length: 447
Score: 670.00 Matches: 131
Percent Similarity: 99.24% Conservative: 0
Best Local Similarity: 99.24% Mismatches: 1
Query Match: 99.26% Indels: 0
DB: 6 Gaps: 0

US-09-684-215A-4 (1-132) x BD006325 (1-447)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 11 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGTGGCGAGGATTCGCCATTCGGATC 70

QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40
Db 71 GGGCAGGGGATGGCGATCGCGGCCAGATCCGATCGGTGGGGGTCAACCCACCGTTTCAT 130

QY 41 IleGlyProThrAlaPheLeuGlyValValAspAsnGlyAsnGlyAlaArg 60
Db 131 ATCGGCGCTTACCGCTTCTCGGCTGGGTGTTCGACAAACACGCGGCGCAGCA 190

QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
Db 191 GTCCAAACGGTGGTGGGAGCGCTCCGCGGCAAGTCTCGGATCTCCACCGCGCAGTG 250

QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
Db 251 ATCAGCGGGTGGAGCGGCTCCGATCACTCGGCAACCGCGATCGCGGCGCAGCGT 310

QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyThrArg 120
Db 311 GGGCATCATCCCGGTGACGTCTATCTCGTGAACCTGGCAACCAACGATCGGGCGGCGCGGT 370

QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
Db 371 ACAGGGAAGTGCATTTGGCCGAGGACCCCGGCC 406

RESULT 13
BD006445

```



```

LOCUS      BD006445               447 bp    DNA          linear    PAT 31-JAN-2002
DEFINITION Compounds and methods for immunotherapy and diagnosis of
Tuberculosis.
ACCESSION  BD006445
VERSION    BD006445.1 GI:18634816
KEYWORDS   JP 2001501832-A/4.
SOURCE     unidentified
ORGANISM   unclassified.
REFERENCE  1 (bases 1 to 447)
AUTHORS    Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Neto,A.C., Houghton,R.,
            Vedvick,T.H., Twardzik,D.R. and Lodes,M.J.
TITLE      Compounds and methods for immunotherapy and diagnosis of
JOURNAL    Patent: JP 2001501832-A 4 13-FEB-2001;
CORIXA CORP
COMMENT     OS Unidentified
            PN JP 2001501832-A/4
            PD 13-FEB-2001
            PF 07-OCT-1997 JP 1998518456
            PR 11-OCT-1996 US 08/730510,13-MAR-1997 US 08/818112 PI
            STEVEN G REED,YASIR A W SKEIKY,DAVIN C DILLON, PI ANTONIO CAMPOS
            NETO,
            PI RAYMOND HOUGHTON,THOMAS S VEDVICK,DANIEL R TWARDZIK, PI
            MICHAEL J LODES
            PC C12N15/31,C07K14/35,A61K39/04,A61K48/00,A61K49/00,C12N15/62,
            C07K19/00,
            PC G01N33/50,G01N33/60,G01N33/569,C12N1/19,C12N1/20,C12N1/21, PC
            C12N5/10//
            PC (C12N1/21,C12R1:19)
            CC Strandedness: Single;
            CC Topology: Linear;
            FH Key Location/Qualifiers
            FT source 1..447
            /organism="Unidentified".
FEATURES   . source
            Location/Qualifiers
            1..447
            /organism="unidentified"
            /mol_type="genomic DNA"
            /db_xref="taxon:32644"
BASE COUNT 79 a 146 c 149 g 72 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 3,24e-42 Length: 447
Score: 670.00 Matches: 131
Percent Similarity: 99.24% Conservative: 0
Best Local Similarity: 99.24% Mismatches: 1
Query Match: 99.26% Indels: 0
DB: 6 Gaps: 0
US-09-684-215A-4 (1-132) x BD006445 (1-447)
Qy 1 ThrAlaAlaSerAspAsnPhGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 11 ACGGCGCGTCCGATACCTCCAGCTGCCAGGTGGCGAGGATTCGCATTCGCATC 70
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40
Db 71 GGCAGCGCATGCGATCGCGGCCAGATCCGATCGGTGGGGTCCACCCACCGTTTCAT 130
Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArg 60
Db 131 ATCGGGCGCTACCGCCCTCCCTCGGCTGGGTGTTGTCGACAAACGCGGCGACGA 190
Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
Db 191 GTCCACGCGGTGGGAGCGCTCCGCGGCGAAGTCTCGGCATCTCCACCGCGAGGTG 250
Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
Db 251 ATCACCGCGGTGACGCGCTCCGATCACTCGCCACCGCATGGCGGACGCGCTTAC 310
Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120

```

```

Db 311 GGCATCATCCCGTGCATCTCGTGAACCTGGCAACCAAGTCGGCGGCACCGCT 370
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
Db 371 ACAGGGAACGTGACATTTGCCGAGGAGGCCCGCCGCC 406
RESULT 14
LOCUS      BD069285               447 bp    DNA          linear    PAT 27-AUG-2002
DEFINITION Compounds and methods for immunotherapy and diagnosis of
Tuberculosis.
ACCESSION  BD069285
VERSION    BD069285.1 GI:22614888
KEYWORDS   JP 2001517069-A/4.
SOURCE     unidentified
ORGANISM   unclassified.
REFERENCE  1 (bases 1 to 447)
AUTHORS    Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Neto,A.C., Houghton,R.,
            Vedvick,T.H. and Twardzik,D.R.
TITLE      Compounds and methods for immunotherapy and diagnosis of
JOURNAL    Patent: JP 2001517069-A 4 02-OCT-2001;
CORIXA CORP
COMMENT     OS Unidentified
            PN JP 2001517069-A/4
            PD 02-OCT-2001
            PF 30-AUG-1996 JP 1997511464
            PR 01-SEP-1995 US 08/523436, 22-SEP-1995 US 08/533634 PR
            22-MAR-1996 US 08/620874,05-JUN-1996 US 08/659683 PR
            12-JUL-1996 US 08/680574
            PI STEVEN G REED,YASIR A W SKEIKY,DAVIN C DILLON,ANTONIO CAMPOS
            PI NETO,
            PI RAYMOND HOUGHTON,THOMAS H VEDVICK,DANIEL R TWARDZIK PC
            C12N15/31,C07K14/35,A61K38/16,C12N15/62,G01N33/569,C12Q1/68, PC
            C12N5/10,
            PC C12N1/21//A61K39/04,(C12N1/21,C12R1:19)
            CC Strandedness: Single;
            CC Topology: Linear;
            CC Compounds and methods for immunotherapy and diagnosis of CC
            tuberculosis
            FH Key Location/Qualifiers
            FT source 1..447
            /organism="Unidentified".
FEATURES   . source
            Location/Qualifiers
            1..447
            /organism="unidentified"
            /mol_type="genomic DNA"
            /db_xref="taxon:32644"
BASE COUNT 79 a 146 c 149 g 72 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 3,24e-42 Length: 447
Score: 670.00 Matches: 131
Percent Similarity: 99.24% Conservative: 0
Best Local Similarity: 99.24% Mismatches: 1
Query Match: 99.26% Indels: 0
DB: 6 Gaps: 0
US-09-684-215A-4 (1-132) x BD069285 (1-447)
Qy 1 ThrAlaAlaSerAspAsnPhGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 11 ACGGCGCGTCCGATACCTCCAGCTGCCAGGTGGCGAGGATTCGCATTCGCATC 70
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40
Db 71 GGCAGCGCATGCGATCGCGGCCAGATCCGATCGGTGGGGTCCACCCACCGTTTCAT 130
Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArg 60
Db 131 ATCGGGCGCTACCGCCCTCCCTCGGCTGGGTGTTGTCGACAAACGCGGCGACGA 190

```

QY 61 ValGlnArgValValGlySerAlaProAlaAserLeuGlyIleSerThrGlyAspVal 80  
Db 191 GTCCAAACGGTGGTGGGAGCGCTCCGGCGCAAGTCTCGCATCTCCACCGGACGCTG 250  
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 251 ATCACCGGGTCCGAGCGGCTCCGATCAACTCGGCGACCGGATGGCGAGCGCTTAAC 310  
QY 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
Db 311 GGGCATCATCCGGTGACGTCTCGGTGAACCTGGCAACCAAGTCGGCGGCGACGCGT 370  
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
Db 371 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 406

## RESULT 15

ARI69165  
LOCUS 1872 bp DNA linear PAT 17-DEC-2001  
DEFINITION Sequence 17 from patent US 6290969.  
ACCESSION ARI69165  
VERSION ARI69165.1 GI:17906947  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1872)  
AUTHORS Reed, S.G., Skeiky, Y.A.W., Dillon, D.C., Campos-Neto, A., Houghton, R.,  
Vedvick, T.S. and Twardzik, D.R.  
TITLE Compounds and methods for immunotherapy and diagnosis of  
tuberculosis  
JOURNAL Patent: US 6290969-A 17 18-SEP-2001;  
FEATURES Location/Qualifiers  
Source 1..1872  
BASE COUNT 318 a 616 c 604 g 332 t 2 others  
ORIGIN

## Alignment Scores:

Pred. No.: 1.35e-41 Length: 1872  
Score: 670.00 Matches: 131  
Percent Similarity: 99.24% Conservative: 0  
Best Local Similarity: 99.24% Mismatches: 1  
Query Match: 99.26% Indels: 0  
DB: 6 Gaps: 0

US-09-684-215A-4 (1-132) x ARI69165 (1-1872)

QY 1 ThrAlaAserAspAsnPhcGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20  
Db 758 ACGGCGCGCTCCGATAAATCCAGCTGTCCAGGTTGGCAGGATTCGCCATTCCGATC 817  
QY 21 GlyGlnAlaMetAlaIleAlaGlyClnIleArgSerGlyGlyGlySerProThrValHis 40  
Db 818 GGGCAGCGGATGGCGATCCGGGGCAATCCGATCGGTGGGGTACCCACCGCTTCAT 877  
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAlaArg 60  
Db 878 ATCGGGCTACCGCTTCTCGGCTTGGGTGTGTGTCGACACACAGGCAACGGCGCACGA 937  
QY 61 ValGlnArgValValGlySerAlaProAlaAserLeuGlyIleSerThrGlyAspVal 80  
Db 938 GTCCAAACGGTGGTGGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGGCGACGCTG 997  
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 998 ATCACCGGGTCCGAGCGGCTCCGATCAACTCGGCGACCGGATGGCGAGCGCTTAAC 1057  
QY 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
Db 1058 GGGCATCATCCGGTGACGTCTCGGTGAACCTGGCAACCAAGTCGGCGGCGACGCGT 1117

QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
Db 1118 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 1153

Search completed: September 5, 2003, 10:31:39  
Job time : 2415.22 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 5, 2003, 08:28:59 ; Search time 70.5517 Seconds  
(without alignments)  
825.814 Million cell updates/sec

Title: US-09-684-215A-4

Perfect score: 675

Sequence: 1 TAASDNFQLSQGGGGAIPA.....QTKSGGTRTGNVTLAEGPPA 132

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlx  
-O=/cgn2\_1/USPTO\_spool\_p/US09684215/runat\_05092003\_072203\_593/app\_query.fasta\_1.853  
-DB-Issued\_Patents\_NA -QWTF=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09684215 -ECGN\_1\_145 [runat\_05092003\_072203\_593 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued\_Patents\_NA: \*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq: \*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq: \*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq: \*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq: \*  
5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq: \*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	675	100.0	2287	4	US-09-223-040-1
2	675	100.0	4403765	3	US-09-103-840A-2
3	675	100.0	4411529	3	US-09-103-840A-1
4	670	99.3	447	3	US-08-818-112-4
5	670	99.3	447	4	US-08-818-111-4
6	670	99.3	447	4	US-09-056-556-4
7	670	99.3	447	4	US-09-072-596-4
8	670	99.3	1872	3	US-08-818-112-17
9	670	99.3	1872	4	US-08-818-111-17
10	670	99.3	1872	4	US-09-056-556-17
11	670	99.3	1872	4	US-09-072-596-17
12	632	93.6	882	4	US-09-736-457-1862

13	632	93.6	900	4	US-09-643-597-353	Sequence 353, App
14	632	93.6	900	4	US-09-606-421B-353	Sequence 353, App
15	632	93.6	945	4	US-09-736-457-1861	Sequence 1861, App
16	632	93.6	1012	4	US-09-643-597-351	Sequence 351, App
17	632	93.6	1012	4	US-09-606-421B-351	Sequence 351, App
18	632	93.6	1464	4	US-09-620-412C-348	Sequence 348, App
19	632	93.6	1464	4	US-09-598-419-348	Sequence 348, App
20	632	93.6	1557	4	US-09-620-412C-332	Sequence 332, App
21	632	93.6	1557	4	US-09-598-419-332	Sequence 332, App
22	632	93.6	1578	4	US-09-556-877-188	Sequence 188, App
23	632	93.6	1578	4	US-09-620-412C-188	Sequence 188, App
24	632	93.6	1578	4	US-09-598-419-188	Sequence 188, App
25	632	93.6	1752	4	US-09-620-412C-352	Sequence 352, App
26	632	93.6	1752	4	US-09-598-419-352	Sequence 352, App
27	632	93.6	1758	4	US-09-620-412C-336	Sequence 336, App
28	632	93.6	1758	4	US-09-598-419-336	Sequence 336, App
29	632	93.6	1860	4	US-09-620-412C-308	Sequence 308, App
30	632	93.6	1860	4	US-09-598-419-308	Sequence 308, App
31	632	93.6	1896	4	US-09-620-412C-324	Sequence 324, App
32	632	93.6	1896	4	US-09-598-419-324	Sequence 324, App
33	632	93.6	1941	4	US-09-620-412C-316	Sequence 316, App
34	632	93.6	1941	4	US-09-598-419-316	Sequence 316, App
35	632	93.6	1965	4	US-09-620-412C-340	Sequence 340, App
36	632	93.6	1965	4	US-09-598-419-340	Sequence 340, App
37	632	93.6	2052	4	US-09-620-412C-356	Sequence 356, App
38	632	93.6	2052	4	US-09-598-419-356	Sequence 356, App
39	632	93.6	2076	4	US-09-620-412C-312	Sequence 312, App
40	632	93.6	2076	4	US-09-598-419-312	Sequence 312, App
41	632	93.6	2103	4	US-09-620-412C-344	Sequence 344, App
42	632	93.6	2103	4	US-09-598-419-344	Sequence 344, App
43	632	93.6	2148	4	US-09-620-412C-320	Sequence 320, App
44	632	93.6	2148	4	US-09-598-419-320	Sequence 320, App
45	632	93.6	2148	4	US-09-598-419-320	Sequence 320, App

#### ALIGNMENTS

#### RESULT 1

US-09-223-040-1  
; Sequence 1, Application US/09223040  
; Patent No. 6544522  
; GENERAL INFORMATION:  
; APPLICANT: Skeiky, Yasir  
; APPLICANT: Alderson, Mark  
; APPLICANT: Campos-Neto, Antonio  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens  
; TITLE OF INVENTION: and Their Uses  
; FILE REFERENCE: 014038-0090100S  
; CURRENT APPLICATION NUMBER: US/09/223.040  
; CURRENT FILING DATE: 1998-12-30  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 2287  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: tri-fusion  
; OTHER INFORMATION: protein Ra12-TbH9-Ra35  
; NAME/KEY: modified\_base  
; LOCATION: (30)  
; OTHER INFORMATION: n = g, a, c or t  
; NAME/KEY: modified\_base  
; LOCATION: (33)  
; OTHER INFORMATION: n = g, a, c or t  
; NAME/KEY: CDS  
; LOCATION: (42)..(2231)  
; NAME/KEY: modified\_base  
; LOCATION: (2270)  
; OTHER INFORMATION: n = g, a, c or t  
US-09-223-040-1

## Alignment Scores:

Pred. No.: 1.85e-61 Length: 2287  
Score: 675.00 Matches: 132  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-684-215A-4 (1-132) x US-09-223-040-1 (1-2287)

QY 1 ThrAlaAlaSerAspAsnPhcInLeuSerGlnGlyGlnGlyPheAlaIleProfil 20  
|||||  
Db 63 ACGGCGCGCTCCGATAACTTCCAGCTGTCCAGGGTGGCAGGGATTCGCCATTCCGATC 122  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40  
|||||  
Db 123 GGGCAGGCGATGGCATCGCGGCCAGATCCATCGGTGGGGTCCACCCACCGTTTAT 182  
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArg 60  
|||||  
Db 183 ATCGGCGCTACCGCTTCTCGGCTTGGGTGTGTGTCGACAAACAGCGCAACGGCGCACGA 242  
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
|||||  
Db 243 GTCCACAGCGGTGGTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGGACGTG 302  
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
|||||  
Db 303 ATCAGCGGCTCGACGGCGCTCCGATCACTCGGCCACCGCGATGGCGGACGGCTTAAC 362  
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
|||||  
Db 363 GGGCATCATCCCGGTGACGTCTCGGTGACCTGGCAACCAAGTCGGCGGCGACGGCT 422  
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
|||||  
Db 423 ACAGGGAACGTGACATTGGCCGAGGAGACCCCGGCC 458

## RESULT 2

US-09-103-840A-2

; Sequence 2, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; FILE REFERENCE: 24366-20007.00  
; CURRENT APPLICATION NUMBER: US/09/103,840A  
; CURRENT FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 4403765  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
; FEATURE:  
; OTHER INFORMATION: CDC 1551  
; OTHER INFORMATION: "n" bases at various positions throughout the sequence  
; OTHER INFORMATION: represent a, t, c or g  
US-09-103-840A-2

## Alignment Scores:

Pred. No.: 2.7e-57 Length: 4403765  
Score: 675.00 Matches: 132  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0

US-09-684-215A-4 (1-132) x US-09-103-840A-2 (1-4403765)

QY 1 ThrAlaAlaSerAspAsnPhcInLeuSerGlnGlyGlnGlyPheAlaIleProfil 20  
|||||  
Db 151984 ACGGCGCGCTCCGATAACTTCCAGCTGTCCAGGGTGGCAGGGATTCGCCATTCCGATC 152043  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40  
|||||  
Db 152044 GGGCAGGCGATGGCATCGCGGCCAGATCCATCGGTGGGGTCCACCCACCGTTTAT 152103  
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArg 60  
|||||  
Db 152104 ATCGGCGCTACCGCTTCTCGGCTTGGGTGTGTGTCGACAAACAGCGCAACGGCGCACGA 152163  
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
|||||  
Db 152164 GTCCACAGCGGTGGTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGACGTG 152223  
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
|||||  
Db 152224 ATCAGCGGCTCGACGGCGCTCCGATCACTCGGCCACCGCATGGCGGACGGCTTAAC 152283  
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
|||||  
Db 152284 GGGCATCATCCCGGTGACGTCTCGGTGACCTGGCAACCAAGTCGGCGGCGACGGCT 152343  
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
|||||  
Db 152344 ACAGGGAACGTGACATTGGCCGAGGAGACCCCGGCC 152379

## RESULT 3

US-09-103-840A-1

; Sequence 1, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; FILE REFERENCE: 24366-20007.00  
; CURRENT APPLICATION NUMBER: US/09/103,840A  
; CURRENT FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 4411529  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
; OTHER INFORMATION: H37Rv  
US-09-103-840A-1

## Alignment Scores:

Pred. No.: 2.7e-57 Length: 4411529  
Score: 675.00 Matches: 132  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0

US-09-684-215A-4 (1-132) x US-09-103-840A-1 (1-4411529)

QY 1 ThrAlaAlaSerAspAsnPhcInLeuSerGlnGlyGlnGlyPheAlaIleProfil 20  
|||||  
Db 151815 ACGGCGCGCTCCGATAACTTCCAGCTGTCCAGGGTGGCAGGGATTCGCCATTCCGATC 151874  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40  
|||||  
Db 151875 GGGCAGGCGATGGCATCGCGGCCAGATCCATCGGTGGGGTCCACCCACCGTTTAT 151934  
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArg 60  
|||||  
Db 151935 ATCGGCGCTACCGCTTCTCGGCTTGGGTGTGTGTCGACAAACAGCGCAACGGCGCACGA 151994



```
Alignment Scores:
Pred. No.: 7.8e-62 Length: 447
Score: 670.00 Matches: 131
Percent Similarity: 99.24% Conservative: 0
Best Local Similarity: 99.24% Mismatches: 1
Query Match: 99.26% Indels: 0
DB: 4 Gaps: 0

US-09-684-215A-4 (1-132) x US-08-818-111-4 (1-447)
QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20
DB 11 ACGCCCGCGTCCGATAACTTCCAGCTGTCCAGGTTGGCAGGATTGCCATTCCGATC 70
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
DB 71 GGGCAGGCGATGGCGATCGCGGCCAGATCCGATCGGTGGGGTTCACCCACCGTTTCAT 130
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
DB 131 ATCGGCCCTACCGCTTCTCGGCTTGGGTGTGTGCGACAAACACGCGCACGCGCACGA 190
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
DB 191 GTCCAAACCGGTGGTGGGAGCGCTCCGCGGCAAGTCTCGGCATCTCCACCGCGACGTG 250
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
DB 251 ATCACC CGGTGGCGGCGCTCCGATCACTCGGCCACCGCGATGGCGGACGCGTTAAC 310
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
DB 311 GGGCATCATCCGCTGAGTCTCATCTCGGTGAACCTGGCAACCAAGTCCGGCGGCGACGGT 370
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
DB 371 ACAGGAACGTGACATTGGCGGAGGACCCCGGCC 406

RESULT 6
US-09-684-215A-4 (1-132) x US-09-056-556-4 (1-447)
QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20
DB 11 ACGCCCGCGTCCGATAACTTCCAGCTGTCCAGGTTGGCAGGATTGCCATTCCGATC 70
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
DB 71 GGGCAGGCGATGGCGATCGCGGCCAGATCCGATCGGTGGGGTTCACCCACCGTTTCAT 130
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
DB 131 ATCGGCCCTACCGCTTCTCGGCTTGGGTGTGTGCGACAAACACGCGCACGCGCACGA 190
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
DB 191 GTCCAAACCGGTGGTGGGAGCGCTCCGCGGCAAGTCTCGGCATCTCCACCGCGACGTG 250
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
DB 251 ATCACC CGGTGGCGGCGCTCCGATCACTCGGCCACCGCGATGGCGGACGCGTTAAC 310
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
DB 311 GGGCATCATCCGCTGAGTCTCATCTCGGTGAACCTGGCAACCAAGTCCGGCGGCGACGGT 370
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
DB 371 ACAGGAACGTGACATTGGCGGAGGACCCCGGCC 406

RESULT 7
US-09-072-596-4
; Sequence 4, Application US/09072596
; Patent No. 6458366
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Campos-Neto, Antonia
; APPLICANT: Houghton, Raymond
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Twardzik, Daniel R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Hendrickson, Ronald C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF TUBERCULOSIS
; NUMBER OF SEQUENCES: 350
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/056.556
; FILING DATE: 07-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.457
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; TREATM
```

APPLICATION NUMBER: US/09/072,596  
FILING DATE: 05-MAY-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.417C9  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 447 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-072-596-4

Alignment Scores:  
Pred. No.: 7.8e-62 Length: 447  
Score: 670.00 Matches: 131  
Percent Similarity: 99.24% Conservative: 0  
Best Local Similarity: 99.24% Mismatches: 1  
Query Match: 99.26% Indels: 0  
Gaps: 4  
DB:

US-09-684-215A-4 (1-132) x US-09-072-596-4 (1-447)

Qy 1 ThrAlaAlaSerAspAsnGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20  
Db 11 ACGCCCGCGTCCGATACTCCAGCTGCCAGGTGGCGAGGATTCGCCATTCGCATC 70  
Qy 21 GlyAlaAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
Db 71 GGCAGCGCATGGCGATCGCGGCCACATCCGATCGGTGGGGGTCAACCCACGCTTCAT 130  
Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAlaArg 60  
Db 131 ATCGGGCTACCGCTTCCCTCGGCTTGGTGTTCGACAAACACGCGGCGCACGA 190  
Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 191 GTCCACCGGTGGTGGGAGCGCTCCGCGGCGCAAGTCTCGCATCTCCACCGCGGCGGTG 250  
Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 251 ATCACCGCGTTCGACGCGCTCCGATCACTCGGCCACCGGATGGCGGCGCTTAAC 310  
Qy 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyThrArg 120  
Db 311 GGCATCATCCGGTGACGTCTCTCGGTGAACCTGGCAACCAAGTCGGGCGGCGACGCT 370  
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
Db 371 ACAGGGAACGTGACATTTGGCGGAGGACCCCGGCC 406

## RESULT 8

US-08-818-112-17

Sequence 17, Application US/08818112

Patent No. 6290969

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.

APPLICANT: Skeiky, Yasir A.W.

APPLICANT: Dillon, Davin C.

APPLICANT: Campos-Neto, Antonio

APPLICANT: Houghton, Raymond

APPLICANT: Vedvick, Thomas S.

APPLICANT: Twardzik, Daniel R.

TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY

TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS

NUMBER OF SEQUENCES: 153

CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED AND BERRY LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7052  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: US/08/818,112  
APPLICATION NUMBER: US/08/818,112  
FILING DATE: 13-MAR-1997  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.411C6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1872 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-818-112-17

Alignment Scores:  
Pred. No.: 4.79e-61 Length: 1872  
Score: 670.00 Matches: 131  
Percent Similarity: 99.24% Conservative: 0  
Best Local Similarity: 99.24% Mismatches: 1  
Query Match: 99.26% Indels: 0  
Gaps: 3  
DB:

US-09-684-215A-4 (1-132) x US-08-818-112-17 (1-1872)

Qy 1 ThrAlaAlaSerAspAsnGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20  
Db 758 ACGCCCGCGTCCGATACTCCAGCTGCCAGGTGGCGAGGATTCGCCATTCGCATC 817  
Qy 21 GlyAlaAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
Db 818 GGCAGCGCATGGCGATCGCGGCCAAATCCGATCGGTGGGGGTCAACCCACGCTTCAT 877  
Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAlaArg 60  
Db 878 ATCGGGCTACCGCTTCCCTCGGCTTGGTGTTCGACAAACACGCGGCGCACGA 937  
Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 938 GTCCACCGGTGGTGGGAGCGCTCCGCGGCGCAAGTCTCGCATCTCCACCGGCGAGTG 997  
Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 998 ATCACCGCGTTCGACGCGCTCCGATCACTCGGCCACCGGATGGCGGCGCTTAAC 1057  
Qy 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyThrArg 120  
Db 1058 GGCATCATCCGGTGACGTCTCTCGGTGAACCTGGCAACCAAGTCGGGCGGCGACGCT 1117  
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
Db 1118 ACAGGGAACGTGACATTTGGCGGAGGACCCCGGCC 1153

## RESULT 9

US-08-818-111-17

Sequence 17, Application US/08818111

Patent No. 6338852

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.

APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Dillon, Davin C.  
APPLICANT: Campos-Neto, Antonia  
APPLICANT: Houghton, Raymond  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Twardzik, Daniel R.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF TUBERCULOSIS  
NUMBER OF SEQUENCES: 148  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/818,111  
FILING DATE: 13-MAR-1997  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.417C6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1872 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-818-111-17

Alignment Scores:  
Pred. No.: 4,79e-61 Length: 1872  
Score: 670.00 Matches: 131  
Percent Similarity: 99.24% Conservative: 0  
Best Local Similarity: 99.24% Mismatches: 1  
Query Match: 99.26% Indels: 0  
DB: 4 Gaps: 0

US-09-684-215A-4 (1-132) x US-08-818-111-17 (1-1872)

QY 1 ThrAlaAlaSerAspAsnPhcInLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
DB 758 ACGGCGCGTCCGATTAACCTCCAGGTGTCACAGGTGGCGAGGATTCGCCATTCGGATC 817  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
DB 818 GGGCAGCGCATGGCGATCCGGGCCAATCCATCCGTTGGGGGTTCACCCACCGTTCAT 877  
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArg 60  
DB 878 ATCCGCGCTACCGCTTCCTCGCTGGGTGGTGTTCGACAAACACGCGACGCGACGA 937  
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
DB 938 GTCCACGCGTGGTGGGAGCGCTCCGGCGCAAGTCTCGGATCTCCACCGCGCGACGTG 997  
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
DB 998 ATCCGCGCGTGGTGGGCGCTCCGATCAACTCCGGCACCGCATTCGCCACCGCGCGTAC 1057  
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyThrArg 120  
DB 1058 GGGCATCATCCGCGTACGTCATCTCGGTGTAAGTGGCAACCAAGTCCGGCGCGACGCGT 1117

QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
DB 1118 ACAGGGAACGTGACATTTGGCCGAGGAGCCCGCCGCC 1153

RESULT 10  
US-09-056-556-17  
Sequence 17, Application US/09056556  
Patent No. 6350456  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Dillon, Davin C.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND  
NUMBER OF SEQUENCES: 241  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/056,556  
FILING DATE: 07-APR-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.457  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1872 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-056-556-17

Alignment Scores:  
Pred. No.: 4,79e-61 Length: 1872  
Score: 670.00 Matches: 131  
Percent Similarity: 99.24% Conservative: 0  
Best Local Similarity: 99.24% Mismatches: 1  
Query Match: 99.26% Indels: 0  
DB: 4 Gaps: 0

US-09-684-215A-4 (1-132) x US-09-056-556-17 (1-1872)

QY 1 ThrAlaAlaSerAspAsnPhcInLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
DB 758 ACGGCGCGTCCGATTAACCTCCAGGTGTCACAGGTGGCGAGGATTCGCCATTCGGATC 817  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
DB 818 GGGCAGCGCATGGCGATCCGGGCCAATCCATCCGTTGGGGGTTCACCCACCGTTCAT 877  
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArg 60  
DB 878 ATCCGCGCTACCGCTTCCTCGCTGGGTGGTGTTCGACAAACACGCGACGCGACGA 937  
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
DB 938 GTCCACGCGTGGTGGGAGCGCTCCGGCGCAAGTCTCGGATCTCCACCGCGCGACGTG 997  
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100



```
Db 998 ATCAGCGGCTGACGGCGCTCCGATCAACTCGGCCACCGGATGGCGGACGCGCTTAAC 1057
Qy 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
Db 1058 GGGCATCATCCGGTGACGTATCTCGGTGAACGGCAACCAAGTCGGGCGGACGCGT 1117
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProLa 132
Db 1118 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 1153

RESULT 11
US-09-072-596-17
; Sequence 17, Application US/09072596
; Patent No. 6458366
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Campos-Neto, Antonia
; APPLICANT: Houghton, Raymond
; APPLICANT: Twardzik, Thomas S.
; APPLICANT: Twardzik, Daniel R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Hendrickson, Ronald C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF TUBERCULOSIS
; NUMBER OF SEQUENCES: 350
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/09/072,596
; FILING DATE: 05-MAY-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Makl, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.417C9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1872 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-072-596-17

Alignment Scores:
Pred. No.: 4.79e-61 Length: 1872
Score: 670.00 Matches: 131
Percent Similarity: 99.24% Conservative: 0
Best Local Similarity: 99.24% Mismatches: 1
Query Match: 99.26% Indels: 0
DB: 4 Gaps: 0

US-09-684-215A-4 (1-132) x US-09-072-596-17 (1-1872)
Qy 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 758 ACGGCGCGTCCGATAAATTCAGCTGTCCAGGGTGGCAGGAGTTCGCCATCCGATC 817
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40

US-09-684-215A-4 (1-132) x US-09-736-457-1862 (1-822)
Qy 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 22 ACGGCGCGTCCGATAAATTCAGCTGTCCAGGGTGGCAGGAGTTCGCCATCCGATC 81
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
Db 82 GGGCAGGCGATGGCGATCGCGGCGCAGATCAAG-----CITCCACCGTTCAT 129
Qy 41 IleGlyProThrAlaPheLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
Db 130 ATCGGGCGCTACCGCTTCCTCGGCTGTGTCGACAAACAACGCGCAACGCGCACGA 189
Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
```

```
|||||
Db 190 GTCCAAACGGCTGGGAGCGCTCCGGCGGAAGTCTCGGATCTCCACCGGCGACGTG 249
QY 81 ILeThraAlaValAspGlyAlaProIleAsnSerAlaThraAlaMetAlaAspAlaLeuAsn 100
|||||
Db 250 ATCAGCGCGGTGCGAGCGCTCCGATCAACTCGGCACCGCGATGGCGGCGGCTTAAAC 309
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
|||||
Db 310 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTCGGCGGCGACGCGT 369
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
|||||
Db 370 ACAGGGAACGTGACATTTGGCCGAGGAGACCCCGGCC 405
RESULT 13
US-09-643-597-353
; Sequence 353, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: AND DIAGNOSIS OF LUNG CANCER
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 353
; LENGTH: 900
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-643-597-353
Alignment Scores:
Pred. No.: 1.78e-57 Length: 900
Score: 632.00 Matches: 126
Percent Similarity: 96.21% Conservative: 1
Best Local Similarity: 95.45% Mismatches: 1
Query Match: 93.63% Indels: 4
DB: 4 Gaps: 1
US-09-684-215A-4 (1-132) x US-09-643-597-353 (1-900)
```

```
QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
|||||
Db 22 ACGGCGCGCTCGGATCAACTTCCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCGGATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40
|||||
Db 82 GGGCAGCGCATGGCGATCGCGGCCAGATCAAG-----CTTCCACCGTTCAT 129
QY 41 IleGlyProThraAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
|||||
Db 130 ATCGGCGCTACCGCTTCTCCGCTTGGGTGTGTGTCGACAAACAGCAACGGCGCACGA 189
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
|||||
Db 190 GTCCAAACGGGTGCGGAGCGCTCCGGCGGCAAGTCTCGGATCTCCACCGGCGACGTG 249
QY 81 ILeThraAlaValAspGlyAlaProIleAsnSerAlaThraAlaMetAlaAspAlaLeuAsn 100
|||||
Db 250 ATCAGCGCGGTGCGAGCGCTCCGATCAACTCGGCACCGCGATGGCGGCGGCTTAAAC 309
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
|||||
Db 310 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTCGGCGGCGACGCGT 369
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
|||||
Db 370 ACAGGGAACGTGACATTTGGCCGAGGAGACCCCGGCC 405
```

```
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
|||||
Db 310 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTCGGCGGCGACGCGT 369
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
|||||
Db 370 ACAGGGAACGTGACATTTGGCCGAGGAGACCCCGGCC 405
RESULT 14
US-09-606-421B-353
; Sequence 353, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: AND DIAGNOSIS OF LUNG CANCER
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 353
; LENGTH: 900
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-606-421B-353
```

```
Alignment Scores:
Pred. No.: 1.78e-57 Length: 900
Score: 632.00 Matches: 126
Percent Similarity: 96.21% Conservative: 1
Best Local Similarity: 95.45% Mismatches: 1
Query Match: 93.63% Indels: 4
DB: 4 Gaps: 1
US-09-684-215A-4 (1-132) x US-09-606-421B-353 (1-900)
```

```
QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
|||||
Db 22 ACGGCGCGCTCGGATCAACTTCCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCGGATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40
|||||
Db 82 GGGCAGCGCATGGCGATCGCGGCCAGATCAAG-----CTTCCACCGTTCAT 129
QY 41 IleGlyProThraAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
|||||
Db 130 ATCGGCGCTACCGCTTCTCCGCTTGGGTGTGTGTCGACAAACAGCAACGGCGCACGA 189
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
|||||
Db 190 GTCCAAACGGGTGCGGAGCGCTCCGGCGGCAAGTCTCGGATCTCCACCGGCGACGTG 249
QY 81 ILeThraAlaValAspGlyAlaProIleAsnSerAlaThraAlaMetAlaAspAlaLeuAsn 100
|||||
Db 250 ATCAGCGCGGTGCGAGCGCTCCGATCAACTCGGCACCGCGATGGCGGCGGCTTAAAC 309
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
|||||
Db 310 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTCGGCGGCGACGCGT 369
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
|||||
Db 370 ACAGGGAACGTGACATTTGGCCGAGGAGACCCCGGCC 405
```

RESULT 15

US-09-736-457-1861  
 ; Sequence 1861, Application US/09736457  
 ; Patent No. 6509448  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Tongtong  
 ; APPLICANT: Bangur, Chaitanya S.  
 ; APPLICANT: Lodes, Michael A.  
 ; APPLICANT: Fanger, Gary  
 ; APPLICANT: Vedvick, Tom  
 ; APPLICANT: Carter, Darrick  
 ; APPLICANT: Retter, Marc  
 ; APPLICANT: Mannion, Jane  
 ; APPLICANT: Fan, Liqun  
 ; APPLICANT: Wang, Aijun  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
 ; FILE OF INVENTION: DIAGNOSIS OF LUNG CANCER  
 ; FILE REFERENCE: 210121.478C15  
 ; CURRENT APPLICATION NUMBER: US/09/736.457  
 ; CURRENT FILING DATE: 2000-12-13  
 ; NUMBER OF SEQ ID NOS: 1864  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 1861  
 ; LENGTH: 945  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-736-457-1861

Alignment Scores:  
 Pred. No.: 1.9e-57 Length: 945  
 Score: 632.00 Matches: 126  
 Percent Similarity: 96.21% Conservative: 1  
 Best Local Similarity: 95.45% Mismatches: 1  
 Query Match: 93.63% Indels: 4  
 Db: 4 Gaps: 1

US-09-684-215A-4 (1-132) x US-09-736-457-1861 (1-945)

Qy	1	ThrAlaAspAspAspPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle	20
Db	22	ACGGCCGGTCCGATAAATTCAGCTGCCAGGGTGGGAGGATCCGCATCCGATC	81
Qy	21	GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis	40
Db	82	GGGCAGGCGATGGCGATCGCGGCCAGATCAAG-----CTTCCACCGTTTCAT	129
Qy	41	IleGlyProThrAlaPheLeuGlyValValAspAsnAsnGlyAlaArg	60
Db	130	ATCGGGCCATCCGCTTCTCGGCTTGGTGTGTTCGACAAACAGCGGCGCACGA	189
Qy	61	ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal	80
Db	190	GTCCAAACCGTGGTGGGAGGCTCCGGCGGCAAGTCTCGGATCTCCACCGGCGCGTG	249
Qy	81	IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn	100
Db	250	ATCACCGGGTTCGACGGGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC	309
Qy	101	GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg	120
Db	310	GGGCATCATCCGGTGGTACGTCTCGGTGACCTGGCAACCAAGTCGGGGCGCACGCGT	369
Qy	121	ThrGlyAsnValThrLeuAlaGluGlyProProAla	132
Db	370	ACAGGGAACGTGACATTGGCGGAGGACCCCGGGCC	405

Search completed: September 5, 2003, 12:11:59  
 Job time : 951.552 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 5, 2003, 07:24:34 ; Search time 256.262 Seconds  
(without alignments)  
1390.474 Million cell updates/sec

Title: US-09-684-215A-4

Perfect score: 675

Sequence: 1 TAASDNFQSGGGGFAIPI.....QTKSGGTRTGNVLAEGPPA 132

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-O=/cgn2.1/uspt03.spool\_p/US09684215/runat\_05092003.072201.548/app\_query.fasta\_1.853  
-DB=N\_Geneseq\_19Jun03 -QWFT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09684215 -CGN\_1\_1\_626 -runat\_05092003.072201.548 -NCPU=6 -ICPU=3  
-NO\_WMAP -LARGQUERY -NEG\_SCORE=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N\_Geneseq\_19Jun03.\*  
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*  
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query % Match	Length	DB ID	Description
1	675	100.0	396	23	AA140769 Nucleotide sequenc
2	675	100.0	672	23	AA140772 Nucleotide sequenc
3	675	100.0	702	20	AA20206 Mycobacterium tube
4	675	100.0	702	23	AA140770 Nucleotide sequenc
5	675	100.0	702	24	AA14140 DNA encoding anti
6	675	100.0	1002	24	AA147077 Mycobacterium tube
7	675	100.0	1002	24	AA147078 Mycobacterium tube
8	675	100.0	1002	24	AA147079 Mycobacterium sp.
9	675	100.0	1002	24	AA147080 Mycobacterium spec
10	675	100.0	1068	20	AA147081 Mycobacterium spec
11	675	100.0	1143	20	AA147082 Mycobacterium spec
12	675	100.0	1629	22	AA147083 M. tuberculosis DN
13	675	100.0	1742	23	AA147084 Nucleotide sequenc
14	675	100.0	1871	23	AA147085 Nucleotide sequenc
15	675	100.0	2190	24	AA147086 Mycobacterium sp.
16	675	100.0	2190	24	AA147087 Mycobacterium spec
17	675	100.0	2191	23	AA147088 Nucleotide sequenc
18	675	100.0	2286	24	AA147089 DNA encoding anti
19	675	100.0	2287	20	AA147090 Mycobacterium tube
20	675	100.0	2287	24	AA147091 Mycobacterium sp.
21	675	100.0	2287	24	AA147092 Mycobacterium spec
22	675	100.0	2808	24	AA147100 Mycobacterium sp.
23	675	100.0	4403765	22	AA199683 Mycobacterium tube
24	675	100.0	4411529	22	AA199682 Mycobacterium tube
25	670	99.3	447	18	AA191466 Mycobacterium tube
26	670	99.3	447	18	AA191403 M. tuberculosis im
27	670	99.3	447	19	AA191450 Mycobacterium tube
28	670	99.3	447	19	AA191432 M. tuberculosis an
29	670	99.3	447	20	AA191952 M. tuberculosis re
30	670	99.3	447	20	AA191940 M. tuberculosis DN
31	670	99.3	447	22	AA191940 M. tuberculosis DN
32	670	99.3	447	22	AA191940 M. tuberculosis DN
33	670	99.3	447	22	AA191940 M. tuberculosis DN
34	670	99.3	447	22	AA191940 M. tuberculosis DN
35	670	99.3	447	22	AA191940 M. tuberculosis DN
36	670	99.3	447	22	AA191940 M. tuberculosis DN
37	670	99.3	447	22	AA191940 M. tuberculosis DN
38	670	99.3	447	22	AA191940 M. tuberculosis DN
39	670	99.3	447	22	AA191940 M. tuberculosis DN
40	670	99.3	447	22	AA191940 M. tuberculosis DN
41	670	99.3	447	22	AA191940 M. tuberculosis DN
42	670	99.3	447	22	AA191940 M. tuberculosis DN
43	646	95.7	1067	22	AA191940 M. tuberculosis DN
44	632	93.6	675	22	AA191940 M. tuberculosis DN
45	632	93.6	675	22	AA191940 M. tuberculosis DN

ALIGNMENTS

RESULT 1

AA140769

ID AA140769 standard; DNA; 396 BP.

AC AA140769;

XX

XX

DT 03-OCT-2002 (first entry)

DE Nucleotide sequence encoding Ral2 protein.

XX Ral2; serine protease antigen; MTB32A; Mycobacterium tuberculosis;

KW vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;

KX immunogen; cytokine; gene; ds.

XX Unidentified.

OS

XX

XX

XX

XX

XX

XX

XX

XX

XX



Alignment Scores:		4.96e-56	Length: 672
Pred. No.:	Score:	675.00	Matches: 132
Percent Similarity:		100.00%	Conservative: 0
Best Local Similarity:		100.00%	Mismatches: 0
Query Match:		100.00%	Indels: 0
DB:		23	Gaps: 0
US-09-684-215A-4 (1-132) x AAL40772 (1-672)			
Qy	1	ThraAlaSerAspAsnPhelGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle	20
Db	25	ACGCCCGCGTCCGATACTTCCAGCTGTCAGGGTGGCAGGATTCCGCAATTCGCGATC	84
Qy	21	GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis	40
Db	85	GGCAGCGATGGCGATCGCGGCAGATCGATCGGTGGGGGTCAACCCACCGTTTCAT	144
Qy	41	IleGlyProThrAlaPheLeuGlyLeuValValAspAsnAsnGlyAlaArg	60
Db	145	ATCGGGCTACCGGCTTCCTCGGCTTGGTGTTCGACAAACACGGCGGCACGA	204
Qy	61	ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal	80
Db	205	GTCCACGCGTGTGGGAGCGCTCCGCGGCAAGTCTCGCATCTCCACCGCGGACGTG	264
Qy	81	IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn	100
Db	265	ATACCGCGGTGACGCGCTCCGATCACTCGGCCACCGCATGGCGGACCGCTTAAC	324
Qy	101	GlyHisHisProGlyAspValIleSerValThrTrpGlnThrIlyssSerGlyGlyThrArg	120
Db	325	GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTCGGGCGGCACGCGT	384
Qy	121	ThrGlyAsnValThrLeuAlaGluGlyProProIle	132
Db	385	ACAGGGAACGTGACATTTGGCCGAGGACCCCGGCC	420
RESULT 4			
AA40770	ID	AA40770 standard; DNA; 702 BP.	
XX	XX	AA40770;	
XX	XX	03-OCT-2002 (first entry)	
XX	XX	Nucleotide sequence encoding Ral2-DPPD fusion protein.	
XX	XX	Ral2; serine protease antigen; MTB32A; Mycobacterium tuberculosis;	
XX	XX	vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;	
XX	XX	immunogen; cytokine; gene; ds.	
XX	XX	Chimeric - Unidentified.	
XX	XX		

Alignment Scores:		4.96e-56	Length: 672
Pred. No.:	Score:	675.00	Matches: 132
Percent Similarity:		100.00%	Conservative: 0
Best Local Similarity:		100.00%	Mismatches: 0
Query Match:		100.00%	Indels: 0
DB:		23	Gaps: 0
US-09-684-215A-4 (1-132) x AAL40772 (1-672)			
Qy	1	ThraAlaSerAspAsnPhelGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle	20
Db	25	ACGCCCGCGTCCGATACTTCCAGCTGTCAGGGTGGCAGGATTCCGCAATTCGCGATC	84
Qy	21	GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis	40
Db	85	GGCAGCGATGGCGATCGCGGCAGATCGATCGGTGGGGGTCAACCCACCGTTTCAT	144
Qy	41	IleGlyProThrAlaPheLeuGlyLeuValValAspAsnAsnGlyAlaArg	60
Db	145	ATCGGGCTACCGGCTTCCTCGGCTTGGTGTTCGACAAACACGGCGGCACGA	204
Qy	61	ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal	80
Db	205	GTCCACGCGTGTGGGAGCGCTCCGCGGCAAGTCTCGCATCTCCACCGCGGACGTG	264
Qy	81	IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn	100
Db	265	ATACCGCGGTGACGCGCTCCGATCACTCGGCCACCGCATGGCGGACCGCTTAAC	324
Qy	101	GlyHisHisProGlyAspValIleSerValThrTrpGlnThrIlyssSerGlyGlyThrArg	120
Db	325	GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTCGGGCGGCACGCGT	384
Qy	121	ThrGlyAsnValThrLeuAlaGluGlyProProIle	132
Db	385	ACAGGGAACGTGACATTTGGCCGAGGACCCCGGCC	420
RESULT 3			
AAZ20206	ID	AAZ20206 standard; DNA; 702 BP.	
XX	XX	AAZ20206;	
XX	XX	17-JAN-2000 (first entry)	
XX	XX	Mycobacterium tuberculosis antigen fusion protein Mtb24 DNA.	
XX	XX	Tuberculosis; antigen; fusion protein; Mtb24; Ral2; DPPD;	
XX	XX	diagnosis; therapy; vaccine; immunogen; ss.	
XX	XX	Mycobacterium tuberculosis.	
XX	XX	W09951748-A2.	
XX	XX	14-OCT-1999.	
XX	XX	07-APR-1999; 99WO-US07717.	
XX	XX	07-APR-1998; 98US-0056556.	
XX	XX	30-DEC-1998; 98US-0223040.	
XX	XX	(CORI-) CORIXA CORP.	
XX	XX	Skeiky YAW, Alderson M, Campos-Neto A;	
XX	XX	WPT; 1999-601610/51.	
XX	XX	P-PSDB; AAY32071.	
XX	XX	New fusion proteins useful for diagnosis, prevention and treatment of	
XX	XX	tuberculosis -	

FH Key Location/Qualifiers  
 FT CDS 4..696 /\*tag= a  
 FT /product= "Ra12-DPPD fusion protein"  
 XX  
 PN WO200125401-A2.  
 PD 12-APR-2001.  
 XX 06-OCT-2000; 2000WO-US27652.  
 XX 07-OCT-1999; 99US-0158585.  
 XX (CORI-) CORIXA CORP.  
 PA Skelky Y, Guderian J;  
 PI  
 XX WPI: 2001-266299/27.  
 DR P-PSDB; AAO22139.  
 XX  
 XX Recombinant nucleic acid molecule for producing high yield expression  
 PT of desired fusion polypeptides, encodes fusion polypeptide comprising  
 PT Mycobacterium tuberculosis coding sequence and heterologous polypeptide  
 PT  
 XX  
 PS Example 1; Fig 3; 39pp; English.  
 XX  
 CC The invention relates to a recombinant nucleic acid molecule encoding a  
 CC fusion polypeptide, comprising a polynucleotide sequence of Ra12, a 14  
 CC kDa C-terminal fragment of serine protease antigen MB32A of  
 CC Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.  
 CC The recombinant fusion nucleic acids and polypeptides are useful for  
 CC providing stable and high yield expression of fusion polypeptides of both  
 CC eukaryotic and prokaryotic origin and to encode a protein product for use  
 CC as an antigen for detecting serum antibodies. The presence of serum  
 CC antibodies to M. tuberculosis antigens in an individual indicates that  
 CC the individual is infected with it. The fusion polypeptides are useful as  
 CC sources of proteins for monitoring binding of serum antibodies to fusion  
 CC proteins and as an immunogen to induce and/or enhance immune responses.  
 CC The coding sequences can be ligated with a coding sequence of another  
 CC molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and  
 CC can be used in vivo as a DNA vaccine. This polynucleotide sequence  
 CC represents the DNA encoding the Ra12-DPPD fusion protein.  
 XX  
 XX Sequence 702 BP; 127 A; 229 C; 225 G; 121 T; 0 other;  
 SQ  
 Alignment Scores:  
 Pred. No.: Length: 5,22e-56 702  
 Score: 675.00 Matches: 132  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 23 Gaps: 0  
 US-09-684-215A-4 (1-132) x AAL40770 (1-702)  
 QY 1 ThrAlaIleSerAspAsnPhcInLeuSerGlnGlyGlnGlyPheAlaIleProfile 20  
 DB 25 ACGGCGCGCTCCGATAACTCCAGCTGTCCAGGTGGCAGGAGTCCGCCATTCGATC 84  
 QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
 DB 85 GGGCAGGCGGTGGCGATCGCGGCGCAGATCCGATCGGTGGGGGTCCACCCACCGTTCAT 144  
 QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyValAsnGlyAlaArg 60  
 DB 145 ATCGGCGCTACCGCTTCTCGGTGGGTGTTCGACACACACGCGACGCGCCACGA 204  
 QY 61 ValGlnArgValValGlySerAlaProAlaIleSerLeuGlyIleSerThrGlyAspVal 80  
 DB 205 GTCCAAACGCTGGTGGGAGCGCTCCGCGCGCAAGTCTCGCATCTCCACCGCGCACGTG 264  
 QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100

DB 265 ATCACCGCGTCCGACGGCGCTCCGATCAACTCCGACCGGATGGCGGACGCGGTTAAC 324  
 QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
 DB 325 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTCCGGGGGCGACGGGT 384  
 QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
 DB 385 ACAGGACGTSACATTGGCGGAGGACCCCGGCC 420  
 RESULT 5  
 ABK14140  
 ID ABK14140 standard; DNA; 702 BP.  
 XX  
 AC ABK14140;  
 XX  
 DT 08-MAY-2002 (first entry)  
 DE  
 DE DNA encoding antigenic fusion protein Ra12-DPPD (Mtb24).  
 KW: Fusion protein; tuberculosis; Mycobacterium tuberculosis; gene; ds;  
 KW tuberculosis; immunogen; vaccine; Ra12-DPPD; Mtb24.  
 XX  
 OS Chimeric - Mycobacterium tuberculosis.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..702  
 FT /\*tag= a  
 FT /product= "Mtb24 #1"  
 FT /partial  
 FT /transl\_except= (pos:694..696, aa:Xaa)  
 FT /note= "No start or stop codon. Xaa= In frame stop codon"  
 FT CDS 2..702  
 FT /\*tag= b  
 FT /product= "Mtb24 #2"  
 FT /partial  
 FT /note= "No start or stop codon"  
 FT /transl\_except= (pos:263..265, aa:Xaa)  
 FT /transl\_except= (pos:353..355, aa:Xaa)  
 FT /transl\_except= (pos:395..397, aa:Xaa)  
 FT /transl\_except= (pos:470..472, aa:Xaa)  
 FT /transl\_except= (pos:701..702, aa:Ser)  
 FT /note= "This codon has an apparent 1 nucleotide deletion  
 FT which alters the reading frame."  
 FT Xaa= In frame stop codon"  
 FT CDS 3..701  
 FT /\*tag= c  
 FT /product= "Mtb24 #3"  
 FT /partial  
 FT /note= "No start or stop codon"  
 FT /transl\_except= (pos:1..2, aa:Pro)  
 FT /transl\_except= (pos:39..41, aa:Xaa)  
 FT /transl\_except= (pos:321..323, aa:Xaa)  
 FT /transl\_except= (pos:339..341, aa:Xaa)  
 FT /transl\_except= (pos:450..452, aa:Xaa)  
 FT /transl\_except= (pos:621..623, aa:Xaa)  
 FT /note= "No start or stop codon. Xaa= In frame stop codon"  
 XX  
 US2002009459-A1.  
 XX  
 PD 24-JAN-2002.  
 XX  
 PF 07-APR-1999; 99US-0287849.  
 XX  
 PR 13-MAR-1997; 97US-0818112.  
 PR 01-OCT-1997; 97US-0942578.  
 PR 18-FEB-1998; 98US-0025197.  
 PR 07-APR-1998; 98US-0056556.  
 PR 30-DEC-1998; 98US-0223040.  
 XX  
 PA (REED/) REED S G.  
 PA (SKEI/) SKEIKY Y A.



PA (DILL/) DILLON D C.  
 PA (ALDE/) ALDERSON M.  
 XX (CAMP/) CAMPOS-NETO A.  
 XX  
 XX Reed SG, Skeiky YA, Dillon DC, Alderson M, Campos-Neto A;  
 XX  
 DR WPI: 2002-171134/22.  
 DR P-PSDB: AAU74600, AAU76541, AAU76542.  
 XX  
 XX New fusion proteins of Mycobacterium tuberculosis antigens, useful for  
 PT diagnosing, treating or preventing M. tuberculosis infection,  
 PT particularly as vaccine for treating or preventing tuberculosis  
 XX  
 PS Examples; Fig 13; 62pp; English.  
 XX  
 CC The invention relates to a purified polypeptide which induces an immune  
 CC response of Mycobacterium tuberculosis. Polypeptides of the invention are  
 CC useful for diagnosing, treating or preventing M. tuberculosis infection,  
 CC particularly tuberculosis infection. In particular, the polypeptides are  
 CC useful as a vaccine formulation with an adjuvant to afford long-term  
 CC protection in animals against the development of tuberculosis. The  
 CC protein coding sequence may be used to encode a protein product for use  
 CC as an immunogen to induce and/or enhance an immune response to M.  
 CC tuberculosis. This sequence represents DNA encoding an M. tuberculosis  
 CC fusion protein of the invention. This polynucleotide encodes 3 different  
 CC proteins, each in a different reading frame.  
 XX  
 SQ Sequence 702 BP; 127 A; 229 C; 225 G; 121 T; 0 other;

## Alignment Scores:

Pred. No.: 5,22e-56 Length: 702  
 Score: 675.00 Matches: 132  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 24 Gaps: 0

US-09-684-215A-4 (1-132) x ABK14140 (1-702)

QY 1 ThrAlaAlaSerAspAsnPhelGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
 DB 25 ACGGCGCGTCCGATCACTCCAGCTGTCAGGGTGGGAGGATCCCATTCGATC 84  
 QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
 DB 85 GGCAGCGCATGGCGATCGCGGCCAGATCCGATCGGGTGGGGTCAACCCCGTTCAT 144  
 QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAlaArg 60  
 DB 145 ATCGGGCCCTACCGCCCTCCCTCGGCTTGGGTGTGTCGACAAACACGCGCACCA 204  
 QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
 DB 205 GTCCACGCGTGTCTGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGCGACGTG 264  
 QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
 DB 265 ATCACCGCGTGTGACGCGCTCGATCAACTCGGCCACCGCGATGCGGACGCGCTAAC 324  
 QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrIleSerGlyThrArg 120  
 DB 325 GGCATCATCCCGGTGACGTCAFTCTGGTGACCTGGCAACCAAGTCGGGCGCGCGGT 384  
 QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
 DB 385 ACAGGAAAGCTGACATTGGCCGAGGACCCCGGCC 420

## RESULT 6

AAAD47077  
 ID AAAD47077 standard; DNA: 1002 BP.  
 XX  
 AC AAAD47077;  
 XX

DT 27-JAN-2003 (first entry)  
 XX  
 DE Mycobacterium tuberculosis mature Ra35 antigen encoding DNA.  
 XX  
 KW Vaccine; immunity; diagnostic agent; gene therapy; Ra35 antigen;  
 KW gene; ds.  
 XX  
 OS Mycobacterium tuberculosis.  
 XX  
 XX Key Location/Qualifiers  
 FT CDS 4..996  
 FT /tag= a  
 FT /product= "Ra35 mature antigenic protein"  
 FT /transl\_except= (pos:547..549, aa:Asp)  
 FT /transl\_except= (pos:550..552, aa:Ser)  
 XX  
 PN WO200272792-A2.  
 XX  
 PD 19-SEP-2002.  
 XX  
 PF 13-MAR-2002; 2002WO-US08223.  
 XX  
 PR 13-MAR-2001; 2001US-275837P.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 XX Skeiky Y, Brannon M, Guderian J;  
 XX  
 DR WPI: 2002-759844/82.  
 DR P-PSDB: AA29702.  
 XX  
 XX New recombinant nucleic acid molecule comprising a Leishmania TSA,  
 PT LeIF, M15 or 6H polynucleotide, useful as vaccine to elicit protective  
 PT immunity against pathogenic microorganisms e.g. Leishmania and  
 PT Mycobacterium tuberculosis  
 XX  
 PS Disclosure; Page 79; 155pp; English.  
 XX  
 CC The invention relates to a recombinant nucleic acid molecule encoding a  
 CC fusion polypeptide. The recombinant nucleic acid comprises a heterologous  
 CC polynucleotide sequence encoding an antigen or an antigenic fragment from  
 CC Mycobacterium sp. and a Leishmania polynucleotide sequence encoding a  
 CC polypeptide or its fragment. The Leishmania polynucleotide is selected  
 CC from TSA, LeIF, M15, and 6H polynucleotides. Sequences of the invention  
 CC are used in methods for eliciting immune response in mammals. They are  
 CC useful as vaccines to elicit protective immunity against pathogenic  
 CC microorganisms such as Leishmania and Mycobacterium tuberculosis. Fusion  
 CC polypeptides are used for enhancing the expression of polynucleotides,  
 CC as in vivo diagnostic agents and for raising antibodies in a non-human  
 CC animal. The invention is used in gene therapy. The present sequence is  
 CC M. tuberculosis mature Ra35 (N-terminus of MTB32A; RA35FL) antigen  
 CC encoding DNA.  
 XX  
 SQ Sequence 1002 BP; 169 A; 331 C; 342 G; 160 T; 0 other;

## Alignment Scores:

Pred. No.: 7.89e-56 Length: 1002  
 Score: 675.00 Matches: 132  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 24 Gaps: 0

US-09-684-215A-4 (1-132) x AAD47077 (1-1002)

QY 1 ThrAlaAlaSerAspAsnPhelGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
 DB 598 ACGGCGCGTCCGATCACTCCAGGTGTCAGGGTGGGAGGATCCCATTCGATC 657  
 QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
 DB 658 GGCAGCGCATGGCGATCGGGGCCAGATCCGATCGGGTGGGGTCAACCCCGTTCAT 717

```
OY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
DB 718 ATCGGGCTACCGCTCTCTCGGCTGGGTGTGTCGACAAACACGACGCGGCACGA 777
OY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
DB 778 GTCCAAACGGTGTGGGAGCGCTCCGGCGCAAGCTCGGCATCTCCACGGCGACGTG 837
OY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
DB 838 ATCACCAGCGGTGCGACGGCGCTCCGATCACTCCGCCACCGCATCGCGACGCGTTAAC 897
OY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
DB 898 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTCGGGGGCGACGGT 957
OY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
DB 958 ACAGGAACGTGACATTCGGCGAGGACCCCGGCC 993
RESULT 7
AAD47078
ID AAD47078 standard; DNA; 1002 BP.
XX
AC AAD47078;
XX
XX
XX 27-JAN-2003 (first entry)
XX
DE Mycobacterium tuberculosis Ra35FLMutSA mutant antigenic protein DNA.
XX
XX Vaccine; immunity; diagnostic agent; gene therapy; Ra35FLMutSA; mutant;
KW gene; antigen; ds.
XX
OS Mycobacterium tuberculosis.
OS Synthetic.
FH Key Location/Qualifiers
FT CDS 4..996
FT FT /*tag= a
FT FT /product= "Ra35FLMutSA mutant antigenic protein"
XX
XX WO200272792-A2.
XX
XX 19-SEP-2002.
XX
XX 13-MAR-2002; 2002WO-US08223.
XX
XX 13-MAR-2001; 2001US-275837P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Skeiky Y, Brannon M, Guderian J;
XX
XX WPI: 2002-759844/82.
XX
XX P-PSDB; AAE29703.
XX
XX New recombinant nucleic acid molecule comprising a Leishmania TSA,
PT LeIF, M15 or 6H polynucleotide, useful as vaccine to elicit protective
PT immunity against pathogenic microorganisms e.g. Leishmania and
PT Mycobacterium tuberculosis
XX
XX Disclosure; Page 80-81; 155pp; English.
XX
XX The invention relates to a recombinant nucleic acid molecule encoding a
CC fusion polypeptide. The recombinant nucleic acid comprises a heterologous
CC polynucleotide sequence encoding an antigen or an antigenic fragment from
CC Mycobacterium sp. and a Leishmania polynucleotide sequence encoding a
CC polypeptide or its fragment. The Leishmania polynucleotide is selected
CC from TSA, LeIF, M15, and 6H polynucleotides. Sequences of the invention
CC are used in methods for eliciting immune response in mammals. They are
CC useful as vaccines to elicit protective immunity against pathogenic
CC microorganisms such as Leishmania and Mycobacterium tuberculosis. Fusion
CC polypeptides are used for enhancing the expression of polynucleotides,
```

```
CC as in vivo diagnostic agents and for raising antibodies in a non-human
CC animal. The invention is used in gene therapy. The present sequence is
CC M. tuberculosis Ra35FLMutSA mutant antigenic protein encoding DNA.
XX
SQ Sequence 1002 BP; 169 A; 331 C; 341 G; 161 T; 0 other;
Alignment Scores:
Pred. No.: 7,89e-56 Length: 1002
Score: 675.00 Matches: 132
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0
US-09-684-215a-4 (1-132) x AAD47078 (1-1002)
OY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
DB 598 ACGGCCGGTCCGATAAATTCAGCTGTCCAGGGTGGCAGGAGGATTCGCCATTCGGATC 657
OY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
DB 658 GGGCAGCGGATGGCGATCGCGGCCAGATCCGATCGGTGGGGGTCTACCCACCGCTTCA 717
OY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
DB 718 ATCGGGCTACCGCTCTCTCGGTGTTGTGTCGACAAACACGACGCGGCACGA 777
OY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
DB 778 GTCCAAACGGTGTGGGAGCGCTCCGGCGCAAGCTCGGCATCTCCACGGCGACGTG 837
OY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
DB 838 ATCACCAGCGGTGCGACGGCGCTCCGATCACTCCGCCACCGCATCGCGACGCGTTAAC 897
OY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
DB 898 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTCGGGGGCGACGGT 957
OY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
DB 958 ACAGGAACGTGACATTCGGCGAGGACCCCGGCC 993
RESULT 8
AAD28336
ID AAD28336 standard; cDNA; 1002 BP.
XX
AC AAD28336;
XX
XX 22-APR-2002 (first entry)
XX
DE Mycobacterium sp. MTB32A (Ra35FL) mature protein encoding cDNA.
XX
XX Fusion protein; antigen; serological sensitivity; immune response;
KW tuberculosis; infection; vaccine; MTB32A; Ra32FL; ss.
XX
XX Mycobacterium sp.
XX
XX Key Location/Qualifiers
XX CDS 4..996
XX FT /*tag= a
XX FT /product= "Ra35 mature protein"
XX FT /transl_except= (pos:547..549, aa:Asp)
XX FT /transl_except= (pos:550..552, aa:Ser)
XX
XX WO200198460-A2.
XX
XX 27-DEC-2001.
XX
XX 20-JUN-2001; 2001WO-US19959.
XX
XX 20-JUN-2000; 2000US-0597796.
PR
```

```
PR 01-FEB-2001; 2001US-265737P.
XX (CORI-) CORIXA CORP.
XX Skeiky Y, Reed S, Alderson M;
XX WPI; 2002-147798/19.
DR P-PSDB; AAE17566.
XX Composition comprising MTB39 antigen and MTB32A antigen from
PT Mycobacterium species, useful for eliciting immune response in a
PT subject
XX
PS Disclosure; Page 95; 136pp; English.
XX
CC The present invention relates to fusion proteins containing at least
CC two Mycobacterium species antigens, nucleotides encoding them and
CC compositions comprising such fusion proteins. The present invention
CC particularly relates to nucleic acids encoding fusion proteins that
CC include two or more individual M. tuberculosis antigens which increase
CC the serological sensitivity of sera from individuals infected with
CC tuberculosis and methods for their use in diagnosis, prevention and
CC treatment of tuberculosis infection. Sequences of the invention are
CC useful for eliciting an immune response in a mammal, e.g., human,
CC immunised with BCG. They are useful in the diagnosis, treatment and
CC prevention of Mycobacterium infection. The fusion proteins and the
CC polynucleotides are useful as diagnostic tools in patients infected
CC with Mycobacterium, in vitro and in vivo assays for detecting humoral
CC antibodies or cell-mediated immunity against M. tuberculosis, for the
CC diagnosis of an infection or monitoring of disease progression, as
CC immunogens to generate or elicit a protective immune response in a
CC patient and for raising anti-M. tuberculosis antibodies in a non-human
CC animal. Sequences of the invention are also used as vaccines. MTB32A
CC fusion proteins of the invention are useful as in vivo diagnostic agents
CC for intradermal skin test. The present sequence is a cDNA encoding
CC Mycobacterium species MTB32A (Ra32FL) mature protein.
XX
SQ Sequence 1002 BP; 169 A; 331 C; 342 G; 160 T; 0 other;
Alignment Scores:
Pred. No.: 7.89e-56 Length: 1002
Score: 675.00 Matches: 132
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0
US-09-684-215A-4 (1-132) x AAD28336 (1-1002)
QY 1 ThrAlaAlaSerAspAsnPhGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
DB 598 ACGGCGCGTCCGATTAACCTCCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCCGATC 657
QY 21 GlyClnAlaMetAlaIleAlaGlnIleArgSerGlyGlyGlySerProThrValHis 40
DB 658 GGGCAGCGATGCGATGCGGGCCAGATCCGATCGGGTGGGGGTCCACCCACGTTTCAT 717
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyValAlaArg 60
DB 718 ATCGGGCTACCGCTTCCTCGCTGGGTGGGTGTCGACAAACACGCGACGGCGACGA 777
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
DB 778 GTCCACGCGTGGTCCGGAGCGCTCCGGCGCAAGTCTCGCATCTCCACCGCGGCGGTG 837
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
DB 838 ATCACCAGCGTCCAGCGCGCTCCGATCAACTCGGCCACCGCGATGCGCGGCTTAAC 897
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpClnThrLysSerGlyGlyThrArg 120
DB 898 GGGCATCATCCCGGTGACGTCTCTCGTGACCTGGCAACCAAGTGGCGCGGCGCACGCT 957
```

```
121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
|||||
958 ACAGGGAACGTGACATTGGCCGAGGACCCCGGCC 993

RESULT 9
AAD28337
ID AAD28337 standard; cDNA; 1002 BP.
XX
AC AAD28337;
XX
DT 22-APR-2002 (first entry)
XX
DE Mycobacterium species Ra35FLMutSA mutant cDNA.
XX
KW Fusion protein; antigen; serological sensitivity; immune response;
KW tuberculosis; infection; vaccine; Ra32FLMutSA; mutant; mutein; ss.
XX
OS Mycobacterium sp.
XX
FH Key Location/Qualifiers
FT CDS 4..996
FT /tag= a
FT /product= "Ra35FLMutSA protein"
XX
PN WO200198460-A2.
XX
XX 27-DEC-2001.
XX
XX 20-JUN-2001; 2001WO-US19959.
XX
XX 20-JUN-2000; 2000US-0597796.
XX
XX 01-FEB-2001; 2001US-265737P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Skeiky Y, Reed S, Alderson M;
XX
XX WPI; 2002-147798/19.
XX
XX P-PSDB; AAE17567.
XX
XX Composition comprising MTB39 antigen and MTB32A antigen from
XX Mycobacterium species, useful for eliciting immune response in a
XX subject
XX
XX Claim 67; Page 96-97; 136pp; English.
XX
CC The present invention relates to fusion proteins containing at least
CC two Mycobacterium species antigens, nucleotides encoding them and
CC compositions comprising such fusion proteins. The present invention
CC particularly relates to nucleic acids encoding fusion proteins that
CC include two or more individual M. tuberculosis antigens which increase
CC the serological sensitivity of sera from individuals infected with
CC tuberculosis and methods for their use in diagnosis, prevention and
CC treatment of tuberculosis infection. Sequences of the invention are
CC useful for eliciting an immune response in a mammal, e.g., human,
CC immunised with BCG. They are useful in the diagnosis, treatment and
CC prevention of Mycobacterium infection. The fusion proteins and the
CC polynucleotides are useful as diagnostic tools in patients infected
CC with Mycobacterium, in vitro and in vivo assays for detecting humoral
CC antibodies or cell-mediated immunity against M. tuberculosis, for the
CC diagnosis of an infection or monitoring of disease progression, as
CC immunogens to generate or elicit a protective immune response in a
CC patient and for raising anti-M. tuberculosis antibodies in a non-human
CC animal. Sequences of the invention are also used as vaccines. MTB32A
CC fusion proteins of the invention are useful as in vivo diagnostic agents
CC for intradermal skin test. The present sequence is a cDNA encoding
CC Mycobacterium species Ra32FL mature protein mutant, Ra32FLMutSA.
XX
XX Sequence 1002 BP; 169 A; 331 C; 341 G; 161 T; 0 other;
Alignment Scores:
Pred. No.: 7.89e-56 Length: 1002
Score: 675.00 Matches: 132
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0
```



PI Gicquel B, Lim EM, Pelicic V, Portnoi D, Goguet de la Salmoniere Y;  
 PI Guigueno A;  
 XX WPI; 1999-181045/15.  
 DR P-PSDB; AAY04830.  
 XX Mycobacterial DNA vectors containing reporter constructs - for  
 PT identifying coding or promoter sequences involved in  
 PT infection-associated protein expression  
 XX Claim 22; Fig 50F; 309pp; French.  
 XX Sequences AAX34001-X34252 represent nucleic acids encoding secreted  
 CC proteins from various Mycobacterium species-microorganisms. The  
 CC nucleotide sequences can be used as primers and probes for methods  
 CC for detecting and identifying mycobacteria, especially belonging to  
 CC the M. tuberculosis complex. The encoded proteins can be used in  
 CC vaccines for immunisation against a bacterial or viral infection.  
 XX Sequence 1143 BP; 189 A; 373 C; 395 G; 186 T; 0 other;  
 SQ Alignment Scores:  
 Pred. No.: 9.2e-56 Length: 1143  
 Score: 675.00 Matches: 132  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 20 Gaps: 0

US-09-684-215A-4 (1-132) x AAX34252 (1-1143)

Qy 1 ThrAlaLaSerAspAsnPhleGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
 Db 745 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGTTGGCGAGGATTCGCATCCGCATC 804

Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
 Db 805 GGGCAGGCGATGGCGATGGCGGCGCAGATCGGTTGGGGTGGGGGTCCACCGTTTCAT 864

Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyValAlaArg 60  
 Db 865 ATCGGGCTACCGCTTCTCGGCTTGGTGTTCGACACACACGCGGCGCAGCA 924

Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
 Db 925 GTCCAACGCGTGTGGCGAGCGCTCCGCGGCAAGTCTCGCATCTCCACCGGCGAGTG 984

Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
 Db 985 ATCACCAGGCTGCGAGCGCTCCGATCAACTCGGCCACCGGATGGCGGAGCGCTTAAC 1044

Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
 Db 1045 GGGCATCATCCGCTGAGCTCATCTCGTGNACCTGGCAACCAAGTCCGGCGGCGCGGT 1104

Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
 Db 1105 ACAGGGAACGTGACATTGGCGGAGGAGGCCCGCCGCC 1140

RESULT 12  
 AAS03793  
 ID AAS03793 standard; DNA; 1629 BP.  
 XX AAS03793;  
 AC AAS03793;  
 XX 29-AUG-2001 (first entry)  
 XX M. tuberculosis DNA encoding antigen HTCC#1 fusion protein #4.  
 DE TBRA12-HTCC#1; antigen; vaccine; tuberculosis;  
 XX AIDS; acquired immunodeficiency disease; His Tag; ds.  
 KW Mycobacterium tuberculosis.  
 OS

## Synthetic.

OS Key Location/Qualifiers  
 XX CDS 1..1629  
 FT /\*tag= a  
 FT /product= "TBRA12-HTCC#1"  
 FT /transl\_except= (pos:1621..1623,aa:Xaa)  
 FT /note= "Xaa= In frame STOP codon"  
 FT /partial  
 FT /note= "No start or stop codon. Although the sequence  
 FT does contain an in frame stop codon, 2 further amino  
 FT acids are shown in Figure 8 as being encoded by the  
 FT present sequence, without a further stop codon."  
 FT 25..426  
 FT misc\_feature  
 FT /\*tag= b  
 FT /note= "Region derived from TBRA12"  
 FT 427..444  
 FT misc\_feature  
 FT /\*tag= c  
 FT /note= "Region derived from Thrombin"  
 FT 445..1629  
 FT misc\_feature  
 FT /\*tag= d  
 FT /note= "Region derived from HTCC#1"

WO200124820-A1.

12-APR-2001.

10-OCT-2000; 2000WO-US28095.

07-OCT-1999; 99US-0158338.

07-OCT-1999; 99US-0158425.

(CORI-) CORIXA CORP.

Skeiky Y, Reed S, Houghton RL, McNeill PD, Dillon DC, Lodes ML;

WPI; 2001-290576/30.

Vaccinating against Mycobacteria infections in mammals using fusion  
 proteins comprising combinations of heterologous antigens

Example 2; Fig 8; 168pp; English.

The sequence encodes Mycobacterium tuberculosis fusion protein,  
 TBRA12-HTCC#1 and includes a His tag at the N-terminus to aid  
 purification. Compositions comprising at least 2 heterologous antigens,  
 as a fusion protein, and vectors expressing the fusion proteins are  
 used as vaccines to prophylactically immunise mammals (especially  
 humans) against infection by Mycobacteria. The compositions  
 contain at least 2 heterologous antigens that increase the serological  
 sensitivity of individuals infected with tuberculosis, a disease  
 frequently affecting patients with acquired immunodeficiency disease,  
 AIDS.

SQ Sequence 1629 BP; 298 A; 500 C; 535 G; 296 T; 0 other;

## Alignment Scores:

Pred. No.: 1.39e-55 Length: 1629  
 Score: 675.00 Matches: 132  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 22 Gaps: 0

US-09-684-215A-4 (1-132) x AAS03793 (1-1629)

Qy 1 ThrAlaLaSerAspAsnPhleGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
 Db 25 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGTTGGCGAGGATTCGCATCCGCATC 84

Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
 Db 85 GGGCAGGCGATGGCGATGGCGGCGCAGATCCGATCGGGTGGGGGTCCACCGCTTCAT 144

```
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArg 60
|
|
|
Db 145 ATCGGCGCTACCGCTTCCTCGGCTGGGTGTGTCGACAAACAGCGCAACGGCGCACGA 204
|
|
|
QY 61 ValGlnArgValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
|
|
|
Db 205 GTCCAAACGGGTGGTGGGAGCGCTCGGCGGCAAGTCTCGGCATCTCCACCGCGACGTG 264
|
|
|
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
|
|
|
Db 265 ATCACCAGCGGTGACGCGCTCGGATCACTCGGCACCGCGATGGCGGACGCGCTTAAC 324
|
|
|
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
|
|
|
Db 325 GGGCATCATCCCGTGACGTATCTCGGTGACGTATCTCGGTGACCTGGCAACCAAGTCGGCGGCGACGCT 384
|
|
|
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
|
|
|
Db 385 ACAGGGAACGTGACATTTGGCCGAGGACCCCGGCC 420
|
|
|
RESULT 13
AAL40771
ID AAL40771 standard; DNA; 1742 BP.
XX
AC AAL40771;
XX
DT 03-OCT-2002 (first entry)
XX
DE Nucleotide sequence encoding Ral2-WT1 fusion protein.
XX
KW Ral2; serine protease antigen; MTB32A; Mycobacterium tuberculosis;
KW vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;
KW immunogen; cytokine; gene; ds.
XX
OS Chimeric - Unidentified.
XX
FH Key Location/Qualifiers
FT CDS 4..1740
FT FT /*tag= a
FT FT /product= "Ral2-WT1 fusion protein"
XX
PN WO200125401-A2.
XX
PD 12-APR-2001.
XX
PF 06-OCT-2000; 2000WO-US27652.
XX
PR 07-OCT-1999; 99US-0158585.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skelky Y, Guderian J;
XX
DR WPI: 2001-266299/27.
DR P-PSDB; AA022140.
XX
PT Recombinant nucleic acid molecule for producing high yield expression
PT of desired fusion polypeptides, encodes fusion polypeptide comprising
PT Mycobacterium tuberculosis coding sequence and heterologous polypeptide
XX
PS Disclosure; Fig 4; 39pp; English.
XX
CC The invention relates to a recombinant nucleic acid molecule encoding a
CC fusion polypeptide, comprising a polynucleotide sequence of Ral2, a 14
CC kDa C-terminal fragment of serine protease antigen MTB32A of
CC Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.
CC The recombinant fusion nucleic acids and polypeptides are useful for
CC providing stable and high yield expression of fusion polypeptides of both
CC eukaryotic and prokaryotic origin and to encode a protein product for use
CC as an antigen for detecting serum antibodies. The presence of serum
CC antibodies to M. tuberculosis antigens in an individual indicates that
CC
```

the individual is infected with it. The fusion polypeptides are useful as sources of proteins for monitoring binding of serum antibodies to fusion proteins and as an immunogen to induce and/or enhance immune responses. The coding sequences can be ligated with a coding sequence of another molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and can be used in vivo as a DNA vaccine. This polynucleotide sequence represents the DNA encoding the Ral2-WT1 fusion protein.

Sequence 1742 BP; 372 A; 569 C; 490 G; 311 T; 0 other;

Alignment Scores: 1.5e-55 Length: 1742  
Pred. No.: 675.00 Matches: 132  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 23 Gaps: 0  
DB: 23

US-09-684-215A-4 (1-132) x AAL40771 (1-1742)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20  
|  
|  
|  
Db 25 ACAGCGCGGTGCGGATCACTCCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCCGATC 84  
|  
|  
|  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
|  
|  
|  
Db 85 GGGCAGGGGATGGCGATCGCGGCCAGATCCGATCGGTGGGGGTTCACCCACCGTTCAT 144  
|  
|  
|  
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArg 60  
|  
|  
|  
Db 145 ATCGGCGCTACCGCTTCCTCGGCTGGGTGTGTCGACAAACAGCGCAACGGCGCACGA 204  
|  
|  
|  
QY 61 ValGlnArgValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
|  
|  
|  
Db 205 GTCCAAACGGGTGGTGGGAGCGCTCGGCGGCAAGTCTCGGCATCTCCACCGCGACGTG 264  
|  
|  
|  
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
|  
|  
|  
Db 265 ATCACCAGCGGTGACGCGCTCGGATCACTCGGCACCGCGATGGCGGACGCGCTTAAC 324  
|  
|  
|  
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
|  
|  
|  
Db 325 GGGCATCATCCCGTGACGTATCTCGGTGACGTATCTCGGTGACCTGGCAACCAAGTCGGCGGCGACGCT 384  
|  
|  
|  
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
|  
|  
|  
Db 385 ACAGGGAACGTGACATTTGGCCGAGGACCCCGGCC 420  
|  
|  
|  
RESULT 14  
AAL40768  
ID AAL40768 standard; DNA; 1871 BP.  
XX  
AC AAL40768;  
XX  
DT 03-OCT-2002 (first entry)  
XX  
DE Nucleotide sequence encoding MTB32A protein.  
XX  
KW Ral2; serine protease antigen; MTB32A; Mycobacterium tuberculosis;  
KW vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;  
KW immunogen; cytokine; gene; ds.  
XX  
OS Mycobacterium tuberculosis.  
XX  
FH Key Location/Qualifiers  
FT CDS 89..1156  
FT FT /\*tag= a  
FT FT /product= "Bacillus mycobacterium MTB32A protein"  
XX  
PN WO200125401-A2.  
XX  
PD 12-APR-2001.

PF 06-OCT-2000; 2000WO-US27652.  
XX 07-OCT-1999; 99US-0158585.  
XX (CORI-) CORIXA CORP.  
XX Skeiky Y, Guderian J;  
XX WPI; 2001-266299/27.  
DR P-PSDB; AAO22137.  
XX  
XX Recombinant nucleic acid molecule for producing high yield expression  
PT of desired fusion polypeptides, encodes fusion polypeptide comprising  
PT Mycobacterium tuberculosis coding sequence and heterologous polypeptide  
PT  
XX  
XX Disclosure; Fig 1; 39pp; English.  
XX  
XX The invention relates to a recombinant nucleic acid molecule encoding a  
CC fusion polypeptide, comprising a polynucleotide sequence of Ral2, a 14  
CC kDa C-terminal fragment of serine protease antigen MTB32A of  
CC Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.  
CC The recombinant fusion nucleic acids and polypeptides are useful for  
CC providing stable and high yield expression of fusion polypeptides of both  
CC eukaryotic and prokaryotic origin and to encode a protein product for use  
CC as an antigen for detecting serum antibodies. The presence of serum  
CC antibodies to M. tuberculosis antigens in an individual indicates that  
CC the individual is infected with it. The fusion polypeptides are useful as  
CC sources of proteins for monitoring binding of serum antibodies to fusion  
CC proteins and as an immunogen to induce and/or enhance immune responses.  
CC The coding sequences can be ligated with a coding sequence of another  
CC molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and  
CC can be used in vivo as a DNA vaccine. This polynucleotide sequence  
CC represents the DNA encoding the Mycobacterium tuberculosis MTB32A  
CC protein.  
XX  
SQ Sequence 1871 BP; 317 A; 617 C; 605 G; 332 T; 0 other;

Alignment Scores:  
Pred. No.: 1.63e-55 Length: 1871  
Score: 675.00 Matches: 132  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 23 Gaps: 0

US-09-684-215A-4 (1-132) x AAL40768 (1-1871)

QY 1 ThrAlaIaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20  
DB 758 ACGGCGCGTCCGATAAATCCAGCTGTCCCAAGGTGGCAGGGATTCCGCCATTCGGATC 817  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
DB 818 GGGCAGCGCATGGCATCGCGGCCAATCCATCCGTTGGGGGTCCACCCACCGTTTCAAT 877  
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArg 60  
DB 878 ATCGGCGCTACCGCTCTCTCGCTGTGGTGTTCGACAAACGCGCAACGGCGCACGA 937  
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
DB 938 GTCCACCGCGTGGTCGGAAGCGCTCGGCGGCAAGTCTCGGCATCTCCACCGCGCACGTG 997  
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
DB 998 ATCACCGCGTCCGCGCGCTCCGATCACTCGGCACCGCATGGCGGACCGCTTAC 1057  
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
DB 1058 GGGCATCATCCGGTGCACGTATCTCGGTGACCTGGCAAAACCAAGTCGGCGCGCACGCT 1117  
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132

DB 1118 ACAGGGACGTGACATTGGCCGAGGACCCCGGCC 1153  
RESULT 15  
AAD47084  
ID AAD47084 standard: DNA; 2190 BP.  
XX  
XX AAD47084;  
XX  
XX 27-JAN-2003 (first entry)  
XX  
XX Mycobacterium sp. MTB72FMutSA fusion protein encoding DNA.  
XX  
XX Vaccine; immunity; diagnostic agent; gene therapy; TbH9; antigen;  
KW Ra35MutSA; Ra12; MTB72MutSA; chimeric; gene; ds.  
XX  
XX Chimeric - Mycobacterium sp.  
OS Chimeric - Mycobacterium tuberculosis.  
XX  
XX  
FH Key Location/Qualifiers  
FT CDS 1..2190  
FT /tag= a  
FT /product= "MTB72FMutSA fusion protein"  
XX  
XX WO200272792-A2.  
XX  
XX 19-SEP-2002.  
XX  
XX 13-MAR-2002; 2002WO-US08223.  
XX  
XX 13-MAR-2001; 2001US-275837P.  
XX  
XX (CORI-) CORIXA CORP.  
XX  
XX Skeiky Y, Brannon M, Guderian J;  
XX  
XX WPI; 2002-759844/82.  
DR P-PSDB; AAE29709.  
XX  
XX New recombinant nucleic acid molecule comprising a Leishmania TSA,  
PT LeIF, M15 or 6F polynucleotide, useful as vaccine to elicit protective  
PT immunity against pathogenic microorganisms e.g. Leishmania and  
PT Mycobacterium tuberculosis  
XX  
XX Disclosure; Page 92-93; 155pp; English.  
XX  
XX The invention relates to a recombinant nucleic acid molecule encoding a  
CC fusion polypeptide. The recombinant nucleic acid comprises a heterologous  
CC polynucleotide sequence encoding an antigen or an antigenic fragment from  
CC Mycobacterium sp. and a Leishmania polynucleotide sequence encoding a  
CC polypeptide or its fragment. The Leishmania polynucleotide is selected  
CC from TSA, LeIF, M15, and 6F polynucleotides. Sequences of the invention  
CC are used in methods for eliciting immune response in mammals. They are  
CC useful as vaccines to elicit protective immunity against pathogenic  
CC microorganisms such as Leishmania and Mycobacterium tuberculosis. Fusion  
CC polypeptides are used for enhancing the expression of polynucleotides,  
CC as in vivo diagnostic agents and for raising antibodies in a non-human  
CC animal. The invention is used in gene therapy. The present sequence is  
CC a DNA encoding MTB72F fusion protein. This fusion protein comprises  
CC Ra35MutSA mutant protein and Ra12 protein from M. tuberculosis and  
CC TbH9 protein from Mycobacterium sp.  
XX  
SQ Sequence 2190 BP; 340 A; 697 C; 797 G; 356 T; 0 other;

Alignment Scores:  
Pred. No.: 1.96e-55 Length: 2190  
Score: 675.00 Matches: 132  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 24 Gaps: 0

US-09-684-215A-4 (1-132) x AAD47084 (1-2190)





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 5, 2003, 08:27:24 ; Search time 2280.41 Seconds  
(without alignments)  
1406.846 Million cell updates/sec

Title: US-09-684-215A-4

Perfect score: 675

Sequence: 1 TAASDNFQLSQGQGFPAIPI.....OTKSGGTRGNTVLAEGPPA 132

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_B/US09684215/runat\_05092003\_072202\_572/app\_query.fasta\_1.853  
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09684215@cgn.1.1.4575 -runat\_05092003\_072202\_572 -NCPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST: \*  
1: em\_estba: \*  
2: em\_esthum: \*  
3: em\_estin: \*  
4: em\_estmu: \*  
5: em\_estov: \*  
6: em\_estpl: \*  
7: em\_estro: \*  
8: em\_etc: \*  
9: gb\_estl: \*  
10: gb\_estc: \*  
11: gb\_hlc: \*  
12: gb\_est3: \*  
13: gb\_est4: \*  
14: gb\_est5: \*  
15: em\_estfun: \*  
16: em\_estom: \*  
17: em\_gss\_hum: \*  
18: em\_gss\_inv: \*  
19: em\_gss\_pln: \*  
20: em\_gss\_vrt: \*  
21: em\_gss\_fun: \*  
22: em\_gss\_mus: \*  
23: em\_gss\_man: \*  
24: em\_gss\_pro: \*  
25: em\_gss\_rtd: \*  
26: em\_gss\_phg: \*  
27: em\_gss\_vrl: \*  
28: gb\_gss1: \*

29: gb\_gss2: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	145.5	21.6	289	29	U82114	U82114 U82114 orde
c 2	126.5	18.7	1438	28	BH770798	BH770798 LLMGtag54
c 3	122	18.1	959	29	BZ549048	BZ549048 pacs1-60
c 4	120.5	17.9	726	14	CB679186	CB679186 OSJNEF02F
c 5	117.5	17.4	590	14	CB925602	CB925602 ABAL_22.F
c 6	114	16.9	603	28	AZ934428	AZ934428 BJ_Ba000
c 7	113	16.7	758	13	BQ514888	BQ514888 EST622303
c 8	106.5	15.8	1033	29	BZ561390	BZ561390 pacs2-164
c 9	103.5	15.3	719	28	AZ933900	AZ933900 BJ_Ba000
c 10	103.5	15.3	947	29	BZ549047	BZ549047 pacs1-60
c 11	103	15.3	726	28	AO989479	AO989479 Rfc00025
c 12	102	15.1	765	9	AJ558965	AJ558965 AJ558965
c 13	100.5	14.9	859	14	CA480614	CA480614 AGENCOURT
c 14	99	14.7	423	9	AI597611	AI597611 tn15f02.x
c 15	97.5	14.4	499	12	BI350520	BI350520 fr32b05.y
c 16	97.5	14.4	870	14	CA473840	CA473840 AGENCOURT
c 17	97	14.4	543	12	BI721127	BI721127 1031054B1
c 18	96	14.2	546	9	AW285510	AW285510 LG1_241_E
c 19	96	14.2	551	9	AW285527	AW285527 LG1_241_G
c 20	96	14.2	1093	13	BU557763	BU557763 AGENCOURT
c 21	96	14.2	1394	29	BZ576076	BZ576076 msh2_4761
c 22	95.5	14.1	677	12	BI378928	BI378928 BFLG1_000
c 23	95	14.1	909	14	CA975497	CA975497 AGENCOURT
c 24	94.5	14.0	728	12	BJ285991	BJ285991 BJ285991
c 25	94.5	14.0	1029	12	BI457360	BI457360 603185681
c 26	94	13.9	434	28	BHG38292	BHG38292 1008021F0
c 27	94	13.9	672	14	CD047378	CD047378 PSB027xJ
c 28	94	13.9	1321	13	BU509638	BU509638 AGENCOURT
c 29	93.5	13.9	607	29	CC345457	CC345457 OGAQ70TH
c 30	93.5	13.9	900	14	CA471201	CA471201 AGENCOURT
c 31	93	13.8	689	12	BM448307	BM448307 DSA024D06
c 32	93	13.8	951	14	CA474469	CA474469 AGENCOURT
c 33	92.5	13.7	452	12	BM874057	BM874057 laa07a02.
c 34	92.5	13.7	860	14	CA471291	CA471291 AGENCOURT
c 35	92.5	13.7	1258	12	BM474560	BM474560 AGENCOURT
c 36	92	13.6	894	29	BZ578796	BZ578796 msh2_6009
c 37	92	13.6	899	14	CA471516	CA471516 AGENCOURT
c 38	91.5	13.6	513	12	BM036648	BM036648 fu80g01.y
c 39	91.5	13.6	545	10	BE921575	BE921575 EST425344
c 40	91.5	13.6	548	13	BQ112264	BQ112264 EST597840
c 41	91.5	13.6	555	13	BQ479512	BQ479512 faa68g09.
c 42	91.5	13.6	570	14	CB364397	CB364397 2F001-P00
c 43	91.5	13.6	613	9	AW595047	AW595047 fk28e03.y
c 44	91.5	13.6	746	14	CD284690	CD284690 G39171.95
c 45	91.5	13.6	810	14	CA470510	CA470510 AGENCOURT

# ALIGNMENTS

RESULT 1

U82114

LOCUS

DEFINITION U82114 ordered cosmid library Mycobacterium leprae genomic clone  
cosmid L-373; contig 64, genomic survey sequence.

ACCESSION

U82114

VERSION

U82114.1

KEYWORDS

GI:3647212

SOURCE

Mycobacterium leprae

ORGANISM

Mycobacterium leprae

Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.

REFERENCE

1 (bases 1 to 289)



## AUTHORS

Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M., Burns,J.L., Kaul,R. and Olsen,M.V.

## TITLE

Whole-Genome-Sequence variation among multiple isolates of *Pseudomonas aeruginosa* library

## JOURNAL

J. Bacteriol., (2002) In press

## COMMENT

Contact: Chris K. Raymond

Genome Center

University of Washington

Box 352145, Seattle, WA 98105-2145, USA

Tel: 2062216954

Fax: 2066857244

Email: craymond@u.washington.edu

Class: shotgun.

## FEATURES

source

1. .959

Location/Qualifiers

/organism="Pseudomonas aeruginosa"

/mol\_type="genomic DNA"

/strain="1-60"

/db\_xref="taxon:287"

/clone="pacs1-60\_1644"

/clone\_lib="pacs1-60"

/note="clinical isolate 1-60 Whole genomic shotgun library."

157 a 314 c 291 g 195 t 2 others

BASE COUNT

ORIGIN

## Alignment Scores:

Pred. No.: 0.0324 Length: 959  
Score: 122.00 Matches: 40  
Percent Similarity: 47.29% Conservative: 21  
Best Local Similarity: 31.01% Mismatches: 44  
Query Match: 18.07% Indels: 24  
DB: 29 Gaps: 5

US-09-684-215A-4 (1-132) x B2549048 (1-959)

Qy 4 SerAspAsnPheGlnLeuSerGlnGly-----GlyClnGlyPheAlaIleProIleGly 21  
Db TCGCAGATCTTACCGGCTTCCGGCGCTTCATGGCGCTTCCTCGCATTCGGATCGAT 319  
Qy 22 GlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHisIle 41  
Db GTCCGCTGAACGTCCGCCGACCACTGGAAGAAGCCGGC-----AAGGTCACTGCG 268  
Qy 42 GlyProThrAlaPheLeuGlyLeuGlyValValAlaAspAsnAsn----- 55  
Db TGGTGGCGGTGGTATCCAGGAAGTGAACAAGGATCTCGCGAGTCC 217  
Qy 56 -----GlyAsnGlyAlaArgValGlnArgValValGlySerAlaProAla 70  
Db TAGGGCTCGACAAGCGCTCGCGCGCTGCTGGCGCAACTGTGGAAGACGGCGCGCG 157  
Qy 71 AlaSerLeuGlyIleSerThrGlyAspValIleThrAlaValAlaAspGlyAlaProIleAsn 90  
Db GCCAAGGTGGCTCGAGGTGGCGGATGATATAGCTTGAACGCCCACTGCATCAAC 97  
Qy 91 SerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerVal 110  
Db GAGTCCCGCACCTCGCGCACCTGGTGGCAACATGAAGCGCGGCGACAGATCAACCTG 37  
Qy 111 ThrTrpGlnThrLysSerGlyGlyThr 119  
Db -----GACGGGGGATCC 25  
RESULT 4  
CB679186/c 726 bp mRNA linear EST 09-APR-2003  
LOCUS OSJNEF02F19.r OSJNEF Oryza sativa (japonica cultivar-group) cDNA  
DEFINITION clone OSJNEF02F19 3', mRNA sequence.  
ACCESSION CB679186  
VERSION CB679186.1 GI:29682911  
KEYWORDS EST.  
SOURCE Oryza sativa (japonica cultivar-group)

## ORGANISM

Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.

## REFERENCE

1 (bases 1 to 726)

## AUTHORS

Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E., Kudrna,B., Dean,R., Soderlund,C., Wing,R. and Wang,G.

## TITLE

Large-scale identification of ESTs involved in the interaction between rice and Magnaporthe grisea

## JOURNAL

Unpublished

## COMMENT

Contact: Rod Wing

Arizona Genomics Institute

University of Arizona

Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ

85721-0088, USA

Tel: 520 626 3967

Fax: 520 621 9288

Email: http://genome.arizona.edu

PCR Primers

FORWARD: gta aac cga cgg cca gtg

BACKWARD: gga aac agc tat gac cat g

Plate: 02 row: F column: 19

Seq primer: gga aac agc tat gac cat g.

## FEATURES

source

1. .726

/organism="Oryza sativa (japonica cultivar-group)"

/mol\_type="mRNA"

/cultivar="Nipponbare"

/db\_xref="taxon:39947"

/clone="OSJNEF02F19"

/tissue\_type="Leaf"

/dev\_stage="3 week"

/lab\_host="DH10B"

/clone\_lib="OSJNEF"

/note="Vector: pBluescript II KS +; Site\_1: EcoRI; Site\_2:

XhoI; Uninfected Control"

221 a 181 c 151 g 173 t

BASE COUNT

ORIGIN

## Alignment Scores:

Pred. No.: 0.0328 Length: 726.  
Score: 120.50 Matches: 41  
Percent Similarity: 44.03% Conservative: 18  
Best Local Similarity: 30.60% Mismatches: 50  
Query Match: 17.85% Indels: 25  
DB: 14 Gaps: 4

US-09-684-215A-4 (1-132) x CB679186 (1-726)

Qy 10 SerGlnGlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAla----- 27  
Db ACATCTGCTGCTGTTGTTTGGCATCCCGTCATCAACTGTTCTGAAAATAGCTCCTCAG 520  
Qy 28 -----GlyGlnIleArgSerGlyGlySerProThrValHisIleGlyPro 43  
Db TTAATTCAGTTGGAAGTTGCGCGTGGTGAATGTGGAGTTTGTCTCCAGATCCA 460  
Qy 44 ThrAlaPheLeuGlyLeuGlyValValAlaAspAsnGlyAsnGlyAlaArgValGlnArg 63  
Db ATCCCATAT-----CAGCTTAATGTTCCGCACTGGATCTATAATATTGCAG 415  
Qy 64 ValValGlySerAlaProAlaAlaSerLeuGly----- 74  
Db GTTCTGGGGCAGTCTGTCGACCAAGCTGTTCTTCTACCACTAGGGTTTGCT 355  
Qy 75 -----IleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAla 92  
Db GGTACCATTTCTTGTGTGATGATCATTTGTCCTGGCGGCGGTAACCTCAAGAGCAAA 295  
Qy 93 ThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerValThrTrp 112  
Db TCTGATCTATCGAGGTTCTGGATGCTTATGGCGTGGACACAGGTGAGCTTGACAATC 235

```

QY 113 GlnThrLysSerGlyClyThrArgThrGlyAsnValThrLeu 126
Db 234 CAAGA-----GGCGCTGAATCTCTGGAGGTAACCTTG 202

RESULT 5
CB925602
LOCUS CB925602
DEFINITION ABAL_22_F06_b1_A012 Absciscic acid-treated seedlings Sorghum bicolor
cDNA clone ABAL_22_F06_A012 3', mRNA sequence.
ACCESSION CB925602
VERSION CB925602.1 GI:30161873
KEYWORDS EST.
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE 1 (bases 1 to 590)
AUTHORS R.R., Liang, C., Sun, F., Sullivan, R., Shah, M., Buchanan, C.D.,
Eastman, A. and Pratt, L.H.
TITLE An EST database from Sorghum: ABAL-treated seedlings
JOURNAL Unpublished
COMMENT Other ESTs: ABAL_22_F06_g1_A012
Contact: Cordonnier-Pratt MW
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in
the Human Genome Center, University of Tokyo Institute of Medical
Science; plant material and RNA prepared at Texas A & M University;
sequencing done in the Laboratory for Genomics and Bioinformatics,
University of Georgia. Sequence ends have been trimmed to exclude
vector and regions below Phred quality 16. Three-prime sequences
are presented as their reverse complement and have been trimmed to
exclude polyA.
Seq primer: Sug3 (CGACCTGCAGCTCGACACACA)
POLYA=Yes.
FEATURES
Location/Qualifiers
1..590
/organism="Sorghum bicolor"
/mol_type="mRNA"
/db_xref="taxon:4558"
/clone="ABAL_22_F06_A012"
/clone_lib="Absciscic acid-treated seedlings"
/vector="pME18S-FL3; Site_1: XhoI; Site_2: XhoI; The
library was prepared from polyA+ RNA from seedlings grown
in hydroponic culture. After 12 days, medium was
supplemented with 1 mM absciscic acid (ABA), while leaves
were misted with a solution of 1 mM ABA. Roots and leaves
were harvested after 3, 6, 12, and 24 hr and material from
all time points was combined prior to RNA isolation.
Double-stranded cDNA was cloned unidirectionally into
different DraIII sites of the pME18S-FL3 vector (5'-prime
DraIII site is CACTGTGTG, 3'-prime DraIII site is CACCATGTG
). XhoI excises the cDNA insert."
BASE COUNT 136 a 137 c 156 g 161 t
ORIGIN
Alignment Scores:
Pred. No.: 0.0509 Length: 590
Score: 117.50 Matches: 37
Percent Similarity: 44.1% Conservative: 20
Best Local Similarity: 28.6% Mismatches: 43
Query Match: 17.4% Indels: 29
DB: 14 Gaps: 4

US-09-684-215A-4 (1-132) x CB925602 (1-590)

```

```

QY 15 GlyPheAlaIleProIleGlyGlnAlaMetAlaIleAla-----Gly 28
Db 6 GGCTTTGCTATCCCATCATCAACTGTACTTAAATCGCTCCTCAGTTAATTGAGTTCCGA 65
QY 29 GlnIleArgSerGlyGlyGlySerProThrValHisIleGlyProThrAlaPheLeuGly 48
Db 66 AAGTTCGCTGGCTGGCTTGAATGTGGACTTTCGCTCCGGATCCCAATTCATAT----- 119
QY 49 LeuGlyValValAspAsnGlyAsnGlyAlaArgValGlnArgValValGlySerAla 68
Db 120 -----CAGCTTAAATGTTCCGACGAGGACTCTTATACCTTAAGTACCTGGGGCAGT 170
QY 69 ProAlaIleSerLeuGlyIleSer-----Thr 77
Db 171 GCTGCAGCCACAGCAGGCTTCTGCTCCACCGCAGGGGTTTGTGGTAAATATTGTTCTG 230
QY 78 GlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAsp 97
Db 231 GGTGATATCAGCGTTCGACGTGGACGCGCAACCTGTTAAGGGCAAAATCTGACCTGTGTGAGG 290
QY 98 AlaLeuAsnGlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGly 117
Db 291 GTTCTGGATGACTATGGCTCGGAGATCAGGTGACCTTGACA----- 332
QY 118 GlyThrArgThrGlyAsnValThrLeu 126
Db 333 ---ATCCGCGCAGGCTCAGAAACCCCTT 356

```

## RESULT 6

```

A2934428
LOCUS A2934428
DEFINITION BJ_Ba0002108r B. japonicum BAC library Bradyrhizobium japonicum
genomic, genomic survey sequence.
ACCESSION A2934428
VERSION A2934428.1 GI:13776488
KEYWORDS GSS.
SOURCE Bradyrhizobium japonicum
ORGANISM Bradyrhizobium japonicum
Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
Tomkins, J.P., Wood, T.C., Stacey, M.G., Loh, J.T., Judd, A., Golcochea
, J.L., Stacey, G., Sadowsky, M.J. and Wing, R.A.
A marker-denge, sequence-ready map of the Bradyrhizobium japonicum
genome
Genome Res. 11 (8), 1434-1440 (2001)
MEDLINE 21376150
PUBMED 11483585
COMMENT Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Class: BAC ends
High quality sequence stop: 553.
Location/Qualifiers
1..603
/organism="Bradyrhizobium japonicum"
/mol_type="genomic DNA"
/stain="USD110"
/db_xref="taxon:375"
/lab_host="E. coli"
/clone_lib="B. japonicum BAC library"
/noe="Vector: pIndigo536; Site_1: HindIII"
BASE COUNT 109 a 201 c 203 g 90 t
ORIGIN
Alignment Scores:
Pred. No.: 0.116 Length: 603
Score: 114.00 Matches: 36

```

Percent Similarity: 43.44%  
 Best Local Similarity: 29.51%  
 Query Match: 16.89%  
 DB: 28  
 Gaps: 4

Conservative: 17  
 Mismatches: 57  
 Indels: 12  
 Gaps: 4

US-09-684-215A-4 (1-132) x AZ934428 (1-603)

Oy 13 GlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGlnIleArgSer 32  
 Db 245 GGCATCGCGTTCGATCGCGCCCAACACCGTGAAGAGCGTGTGGCCACGCTCAAGGAC 304  
 Oy 33 GlyGly-----GlySerProThrValHisIleGlyProThrAlaPheLeuGly 48  
 Db 305 AAGGTTGGTTCAGCCCGCGTGGATCGCGCTGAGATTTCAGCCGGTGACG----- 355  
 Oy 49 LeuGlyValValAspAsnGly-----AsnGlyAlaArgValGlnArgVal 64  
 Db 356 TCGATATCGCGCAGACCGCTCGCATGAGAGCGCCGAGGGCGCTGGTGGCGGAGCGG 415  
 Oy 65 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 84  
 Db 416 CAGCGAACGGTCCGGCGGCGAAGCGCGCATCGAGTCCGCGCGAGTCACTCGGTC 475  
 Oy 85 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 104  
 Db 476 AACGGCGAATCCGTCAGAGCGCGCGAGCTCGCCGCCACCATCGCGCGCATCGCGCC 535  
 Oy 105 GlyAspValIleSerValThrTrpGlnThrIleSerGlyGlyThrArgThrGlyAsnVal 124  
 Db 536 CTGCGATCGTGAGCTTAACGTCGTCACAAAG---GCCAGGAGCAATGTCGTGAACCTC 592  
 Oy 125 ThrLeu 126  
 Db 593 ACCCTC 598

#### RESULT 7

BQ514888/c BQ514888 758 bp mRNA linear EST 07-MAR-2003  
 LOCUS EST622303 Generation of a set of potato cDNA clones for microarray  
 DEFINITION analyses mixed potato tissues Solanum tuberosum cDNA clone STM1081  
 3' end, mRNA sequence.

ACCESSION BQ514888  
 VERSION BQ514888.1 GI:21373757

KEYWORDS EST.

SOURCE Solanum tuberosum (potato)

ORGANISM Solanum tuberosum

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamids; Solanales; Solanaceae; Solanum.

AUTHORS Buell,C.R., Hart,A., Baker,B., Tanksley,S., Fry,W., Smart,C.,  
 Restrepo,S., Griffiths,H., van der Hoeven,R., Tsai,J. and  
 Karamycheva,S.A.

TITLE Generation of a set of potato cDNA clones for microarray analyses

JOURNAL Unpublished

COMMENT Other\_ESTs: EST622302

CONTACT: Robin Buell

INSTITUTION The Institute for Genomic Research

ADDRESS 9712 Medical Center Dr., Rockville, MD 20850, USA

EMAIL: potato-array@tigr.org

ABSTRACT This clone can be obtained from the University of Arizona Genomics  
 Institute. Orders can be made through URL:  
 http://genome.arizona.edu/orders/  
 Seq primer: T7.

#### FEATURES

Source

Location/Qualifiers

1..758  
 /organism="Solanum tuberosum"  
 /mol\_type="mRNA"  
 /cultiivar="Kennebec or Binjite"  
 /db\_xref="taxon:4113"  
 /clone="STM1081"  
 /tissue\_type="mixed tissues"  
 /lab\_host="SOLR"

/clone\_lib="Generation of a set of potato cDNA clones for  
 microarray analyses mixed potato tissues"  
 /note="Vector: pBluescript SK(-); Site\_1: EcoRI; Site\_2:  
 XhoI; supplier: Combination of untreated and Phytophthora  
 infestans-treated libraries of stolons, leaves, leaflets,  
 axillary buds of stem explants, petioles, germinating eyes  
 tubers, or roots."

BASE COUNT 215 a 181 c 128 g 234 t

ORIGIN

Alignment Scores:  
 Pred. No.: 0.19 Length: 758  
 Score: 113.00 Matches: 34  
 Percent Similarity: 42.98% Conservative: 15  
 Best Local Similarity: 29.82% Mismatches: 47  
 Query Match: 16.74% Indels: 18  
 DB: 13 Gaps: 3

US-09-684-215A-4 (1-132) x BQ514888 (1-758)

Oy 10 SerGlnGlyGlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGln 29  
 Db 592 ACATCAGCAGGTGTGGATTGCAATCCCTTCTCAACTGTGTGAGAGTTGTGCCCCAG 533  
 Oy 30 IleArgSerGlyGly-----GlySerProThrValHisIleGlyProThrAla 45  
 Db 532 TTGATCCAATCTGAAAAAGTTCTTCGTGCTGTTGAATATTGAAATCGCTCCAGACCTG 473  
 Oy 46 PheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArgValGlnArgValVal 65  
 Db 472 -----ATTGCCAACCAACTTAATGTTTGCAGAAATGGAGCACTGGTTTCTGTGTA 422  
 Oy 66 GlySerAlaProAlaAlaSerLeuGly----- 74  
 Db 421 GGAATAGTCTCGCAGCGAAGCGGAGCTTCTCTACTACAGGGGTTTTCAGGAAAT 362  
 Oy 75 IleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThrAla 94  
 Db 361 ATAGTGTGTCGATATATTATTGAAGCAGTGGTGCACAAACCTGTTAGGAGTAAAGCAGAG 302  
 Oy 95 MetAlaAspAlaLeuAsnGlyHisProGlyAspValIle 108  
 Db 301 TTGTATAAGCCCTGGTAACATATAACATAGTGTGAAGTT 260

#### RESULT 8

BZ561390/c BZ561390

LOCUS pacs2-164\_3239.y3 pacs2-164 Pseudomonas aeruginosa genomic clone  
 DEFINITION pacs2-164\_3239, genomic survey sequence.

ACCESSION BZ561390

VERSION BZ561390.1 GI:27181349

KEYWORDS GSS.

SOURCE Pseudomonas aeruginosa

ORGANISM Pseudomonas aeruginosa

REFERENCE Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 Pseudomonadaceae; Pseudomonas.

AUTHORS 1 (bases 1 to 1033)

Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,

Burns,J.L., Kaul,R. and Olsen,M.V.

Whole-Genome-Sequence variation among multiple isolates of

Pseudomonas aeruginosa library

J. Bacteriol., (2002) In press

Contact: Chris K. Raymond

Genome Center

University of Washington

Box 352145, Seattle, WA 98105-2145, USA

Tel: 2062216954

Fax: 2066857244

Email: craymond@u.washington.edu

Class: shotgun

Location/Qualifiers

1..1033

/organism="Pseudomonas aeruginosa"

```

/mol_type="genomic DNA"
/strain="2-164"
/db_xref="taxon:287"
/clone="pacs2-164_3239"
/clone_lib="pacs2-164"
/note="clinical isolate 2-164 Whole genomic shotgun
library."
BASE COUNT 188 a 313 c 292 g 236 t 4 others
ORIGIN

Alignment Scores:
Pred. No.: 1-2 Length: 1033
Score: 106.50 Matches: 41
Percent Similarity: 44.20% Conservative: 20
Best Local Similarity: 29.71% Mismatches: 56
Query Match: 15.78% Indels: 21
DB: 29 Gaps: 4

US-09-684-215A-4 (1-132) x BZ561390 (1-1033)
QY 4 SerAspAsnPheGlnLeuSerGlnGlyGln-----GlyPheAlaIleProIleGly 21
DB 713 TCCACGTTTTCACCGCGTCCGCGCTCAATGGCCCTGCTTCGGCATTCGATCGAT 654
QY 22 GlnAlaMetAlaIle-AlaGlyGlnIleArg-----Se 32
DB 653 GTCGCGCTGACCGTTCGCGACGAGTTGAAGAACCCGCGCAAGGTCACTCGCGGTGCG 594
QY 32 rGlyGlyGlySerProThrValHisIleGlyProThrAlaPheLeu-GlyLeuGlyValV 52
DB 593 TGGCGCTGTGATCCAGGAGTGAACAAGGATCTCGCCGATCTTCGGCCCTC----- 541
QY 52 alAspAsnAsnGlyAsnGlyAlaArgValGlnArgValValGlySerAlaProAlaAlas 72
DB 540 -----GACAAGCCGTCCGCGCGCTGCTGGCGCAACTGTGTGAAGACGCCCGCGGCCA 486
QY 72 erLeuGlyIleSerThrGlyAspValIleThrAlaValaspGlyAlaProIleAsnSerA 92
DB 485 AGGTGCGCTGCGAGTGGCGGATGTATCCTCAGCTGAACGCCGCGATCGATCAACGAGT 426
QY 92 laThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerVal----- 110
DB 425 CCGCGACCTGCCGCGCTGGTGGCAACATGAAGCGCGGCGGCAAGATCACTCGGACG 366
QY 111 -----ThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 124
DB 365 TGATTGCAACGCCGCGCAAGTCCCTGAGCATGCGCGTAGCGAGCCTT 316

RESULT 9
AZ933900 719 bp DNA linear GSS 24-APR-2001
LOCUS BJ_Ba0001B03r B. japonicum BAC library Bradyrhizobium japonicum
DEFINITION genomic, genomic survey sequence.
ACCESSION AZ933900
VERSION AZ933900.1 GI:13775960
KEYWORDS GSS.
SOURCE Bradyrhizobium japonicum
ORGANISM Bradyrhizobium japonicum
Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
Bradyrhizobiaceae; Bradyrhizobium.
REFERENCE 1 (bases 1 to 719)
AUTHORS Tomkins,J.P., Wood,T.C., Stacey,M.G., Loh,J.T., Judd,A., Golcochea
,J.L., Stacey,G., Sadowsky,M.J. and Wing,R.A.
TITLE A marker-dense, sequence-ready map of the Bradyrhizobium japonicum
genome
JOURNAL Genome Res. 11 (8), 1434-1440 (2001)
MEDLINE 21376150
PUBMED 11483585
CONTACT: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288

```

```

Fax: 864 656 4293
Email: rwing@clemson.edu
Class: BAC ends
High quality sequence stop: 671.
FEATURES
source Location/Qualifiers
1..719
/organism="Bradyrhizobium japonicum"
/mol_type="genomic DNA"
/strain="USD110"
/db_xref="taxon:375"
/lab_host="E. coli"
/clone_lib="B. japonicum BAC library"
/note="Vector: pIndigo36; Site_1: HindIII"
BASE COUNT 127 a 245 c 235 g 111 t
ORIGIN

Alignment Scores:
Pred. No.: 1.55 Length: 719
Score: 103.50 Matches: 34
Percent Similarity: 37.86% Conservative: 19
Best Local Similarity: 24.29% Mismatches: 62
Query Match: 15.33% Indels: 25
DB: 28 Gaps: 4

US-09-684-215A-4 (1-132) x AZ933900 (1-719)
QY 13 GlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGlnIleArgSer 32
DB 44 GCATCGGCTTTGCGATCCTTGTCAACATGGAGCGGTGTCGTCGCTCCGCCAAGGC 103
QY 33 GlyGly-----GlySerProThrValHisIleGlyProThr 44
DB 104 GCGCGCAAGCGGTGAAGCTCCTGCTCGCGCGCAAGTTGTCAGCGGTGACGCCGAG 163
QY 45 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 64
DB 164 ATCGCGGAAAGCTCGCTTG-----CCTTCGCGACCGCGCGGTGTCGCGACCGTG 217
QY 65 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 84
DB 218 GTCTCGAAGCGTCCCGCAGCAAGCGCGCTGAAATCCTCGATCTGATCACCAGGATC 277
QY 85 AspGlyAlaProIleAsn-----90
DB 278 GACGCGCACACCTGGATGATCCCAATGCTTCGACTACCGCTTCGCGACGCTCCACTC 337
QY 91 SerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerVal 110
DB 338 GCGGTTCCGCGCAGATCGACGGGCGCGCGCGCAAGCCGCTCAAG---CTCGCGATC 394
QY 111 ThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnValThrLeuAlaGlyPro 130
DB 395 GCGCTGGAGACCGCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 454

RESULT 10
BZ549047 947 bp DNA linear GSS 17-DEC-2002
LOCUS pacs1-60_1644.s1 pacs1-60 Pseudomonas aeruginosa genomic clone
DEFINITION pacs1-60_1644, genomic survey sequence.
ACCESSION BZ549047
VERSION BZ549047.1 GI:27152628
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE 1 (bases 1 to 947)
AUTHORS Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
Burns,J.L., Kaul,R. and Olsen,M.V.
TITLE Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
JOURNAL J. Bacteriol., (2002) In press
CONTACT: Chris K. Raymond

```

Genome Center  
University of Washington  
Box 352145, Seattle, WA 98105-2145, USA  
Tel: 2062216954  
Fax: 2066857244  
Email: [craymond@u.washington.edu](mailto:craymond@u.washington.edu)

MEDLINE  
PUBMED  
COMMENT

```

FEATURES
  source
  location/Qualifiers
    1 947
    /organism="Pseudomonas aeruginosa"
    /mol_type="genomic DNA"
    /strain="1-60"
    /db_xref="taxon:287"
    /clone_lib="pacsi-60"
    /clone="pacsi-60_1644"
    /note="clinical isolate 1-60 whole genomic shotgun

```

Tel: (41) 1225 826779  
 Fax: (41) 1225 826779  
 Email: [bsrricepath.ac.uk](mailto:bsrricepath.ac.uk)  
 This is one of 2,122 random reads from the M13  
 annotation of identified clones (BLASTX, BLASTN,  
 coli K12 genome) please see french-Constant  
 Acids Res.  
 Seq primer: M13 Forward

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

BASE COUNT	ORIGIN	178 a	316 c	259 g	194 t
/clone_lib="pacs1-60"					
/note="clinical isolate 1-60 whole genomic shotgun library."					

```

FEATURES
Class: shotgun.
Location/Qualifiers
1. .726
   /organism="Photorhabdus luminescens"
   /mol_type="genomic DNA"
   /strain="W14"
   /db_xref="taxon:29488"
   /clone="PLG00025"
   /dev_stage="primary phase variant"
   /clone_lib="Photorhabdus luminescens
   library"

```

	Alignment Scores:		
Pred.	No.:	2.14	Length:
Score:	103.50		Matches:     37
Percent Similarity:	44.74%		Conservative:  14
Best Local Similarity:	32.46%		Mismatches:  48
Query Match:	15.33%		Indels:       16
DB:	29		Gaps:          3

```

/strain= W14
/db_xref="taxon:29488"
/clone="PLG00025"
/dev_stage="primary phase variant"
/clone_lib="Photorhabdus luminescens strain W14 M13
library"
/note="Genomic DNA from strain W14 was size selected (1-2
kb) and then cloned into M13 Janus."
208 a 152 c 159 g 202 t 5 others
BASE COUNT

```

US-09-684-215A-4 (1-132) x B2549047 (1-947)

Qy	4	SerAspAsn	PheGlnLeuSerGlnGly-----GlyGlnGlyPheAlaIleProIleGly	21
Db	374	TCGAGATCTT	CACCCGTTCCGGGGCGTTCATGGCCCTGCTTCGCATTCGCATCGAT	315
Qy	22	GlnAlaMet	AlaIleAlaGlyGlnIleArgSer-----	32
Db	314	GTCCGATCAAGCT	CGCGACAGCTTGAAAGAACGCGAAGGTTCAGTCGCGCTGGCTG	255
Qy	33	GlyGlyGlySerPro	ThrValHisIleGlyProThrAlaPheLeuGlyLeuGlyValVal	52
Db	254	GGCGT-GGTGATCC	AGGAAGTGACACAGCATCTCCGCGAGTCTTCGGCCCTC-----	205

Alignment Scores:	
pred. No.:	1.76
Score:	103.00
Percent Similarity:	46.55%
Best Local Similarity:	33.62%
Query Match:	15.26%
DB:	28
Gaps:	2
Indels:	11
Mismatches:	52
Conservative:	15
Matches:	39
Length:	726

Qy	33	GlyGlyGlySerProThrValHisIleGlyProThrAlaPheLeuGlyLeuGlyValVal	52
Db	254	GGCGT-GGTGATCCAGGAAGTGAACAACGAGTCTCGCGAGTCTTCGGCCTC-----	205
Qy	53	AspAsnAsnGlyValAsnGlyAlaArgValGlnArgValValGlySerAlaProAlaAlaSer	72
Db	204	---GACAAGCCGTCGCGCGCGCTGGTGGCGAACTGGTGAAGACGCCCGCGCGCGAAG	148
Qy	73	LeuGlyIleSerThrGlyAspValIleThrAlaValAlaAspGlyAlaProIleAsnSerAla	92
Db	147	GGTGGCGCTGCAGTGGCGCGATGTGATCAACAGCCTGAACGGCCAGTCGATCAACAGAGTCC	88
Qy	93	ThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAsp	106
Db	87	GCGGACCTCCCGACCTGGTGGCAACATGAAGCCGGCGCAC	46

US-09-684-215A-4 (1-132) x A0989479 (1-726)

```
Qy      .6 AsnPheGlnLeuSerGlnGlyGln-----GlyPheAlaIleProIle 20
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      179 TCATTCGATAAAAGTGAAATGGTGAACGCCAGAGGGTTGGGGTTCGCTATTCCGACT 23
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      21 GlyGlnAlaMetAlaIleAlaGlnIleArgSerGlyGlyGlySerProThrValHis 40
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      239 GAACTCGCTACTAAGATTATTGCGAAACATTTATCCGTGACGGACGAGTTATCCGGGGGTTTT 29
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAlaAspAsnAsnGlyAsn----- 57
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      299 ATCGGCATAACATCGCAAGCGTGCCACATATTTCGCTTCTTAACGGCAATATTAAACCAG 35
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      58 -----GlyAlaArgValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIle 75
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      359 CTTCAAGGGGCTACGGAGTATTTCAGGTAACTACAAACGGCGCTTCGCCAAAAGCCGGTATT 41
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      76 SerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMet 95
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      419 AAAGTCGGGTGATATCATTTACCAAGCGTCAATAACAACCCCGTATTCTTCGGGGTGAACA 47
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

RESULT 11  
AQ89479  
LOCUS  
726 bp DNA linear GSS 14-AUG-2000  
DEFINITION *Photobacterium luminescens* strain W14 M13 library  
*Photobacterium luminescens* genomic clone PLG00025, genomic survey  
sequence.

Db	359	CTTCAAGGGCTACGAGTATTTTCAGGTAACTCAAAACGGCCCTGCCCAAAAGCCGGTATT	411
Qy	76	SerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMet	95
Db	419	AAAGTGGGTGATATCATTTACCAAGCGCTCAATAACAAACCCGCTATTCTTCGCGGTGAAAACA	479
Qy	96	AlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerValThr	111
Db	479	ATGGATCAAGTGGCTGA-AATACCCCGCCGACGCTGGGTNCTGTCACT	525

sequence.	
ACCESSION	AQ989479
VERSION	AQ989479.1
KEYWORDS	GI:9648073
SOURCE	GSS.
ORGANISM	Photorhabdus luminescens
	Photorhabdus luminescens
	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
	Enterobacteriaceae; Photorhabdus.
REFERENCE	1 (bases 1 to 726)
AUTHORS	firrench-Constant,R.H., Waterfield,N., Burland,V., Perna,N.T., Daborn,P.J., Bowen,D. and Blattner,F.R.
TITLE	A genomic sample sequence of the entomopathogenic bacterium Photorhabdus luminescens W14: potential implications for virulence
JOURNAL	Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)

Qy	96	AlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerValThr	111
		::::	
Db	479	ATGGATCAAGTGCTGA-AATACCCCGGCAGCGTGGGTNCTGTCACT	525
RESULT	12		
AJ558965			
LOCUS	AJ558965	765 bp	mRNA
DEFINITION	AJ558965	Antirrhinum majus whole plant	linear
	018.1_09.a08,	Antirrhinum majus	EST
		mRNA sequence.	
ACCESSION	AJ558965		
VERSION	AJ558965.1	GI:31661537	
KEYWORDS	EST.		
SOURCE	Antirrhinum majus (snapdragon)		





ACCESSION	AI597611	258	AGCATTTGGGGTTCGTCGGTCTGCCACGATTATG	293
VERSION	AI597611.1	GI:4606659		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS	NCI/NINDS-CGAP <a href="http://www.ncbi.nlm.nih.gov/ncicgap">http://www.ncbi.nlm.nih.gov/ncicgap</a> .			
TITLE	National Cancer Institute / National Institute of Neurological Disorders and Stroke, Brain Tumor Genome Anatomy Project (CGAP/BTGP), Tumor Gene Index			
JOURNAL	Unpublished			
COMMENT	Contact: Robert Strausberg, Ph.D. Email: <a href="mailto:cgaps-r@mail.nih.gov">cgaps-r@mail.nih.gov</a> Tissue procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D., Ph.D.			
FEATURES	CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D. CDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://www.bio.llnl.gov/bbrp/image/image.html">www.bio.llnl.gov/bbrp/image/image.html</a> Seq primer: -400P from Gibco.			
source	Location/Qualifiers			
	1..423			
	/organism="Homo sapiens"			
	/mol_type="mRNA"			
	/db_xref="taxon:9606"			
	/clone="IMAGE:2167707"			
	/tissue_type="anaplastic oligodendroglioma"			
	/lab_host="DH10B"			
	/clone_lib="NCI CGAP_Brn25"			
	/note="Organ: brain; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCGAGTGGAGCGCGCATAGGTTTTTTTTTTTTTTTTTTT T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library is normalized, and was constructed by Bento Soares and M.Fatima Bonaldo."			
BASE COUNT	65 a 111 c 139 g 108 t			
ORIGIN				
Alignment Scores:				
Pred. No.:	2.31	Length:	423	
Score:	99.00	Matches:	35	
Percent Similarity:	44.35%	Conservative:	20	
Best Local Similarity:	28.23%	Mismatches:	43	
Query Match:	14.67%	Indels:	26	
DB:	9	Gaps:	7	
US-09-684-215A-4 (1-132) x AI597611 (1-423)				
Qy	7 PheGlnLeuSerGln-----GlyGlyGlnGlyPheAlaIleProIleGlyGlnAla 23			
Db	39 TTGTAAGGCAAGGCGGCCCCCTGGGGCACAGGACGAGCGCGGTCTTCAGGG 98			
Qy	24 MetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThr---ValHisIleGly 42			
Db	99 TCAGGGCGCATTTGGGCAGATGCGCTGTAGTGGGGGCGCACCTCCAGTCTTTGTGTCA 158			
Qy	43 ProThrAlaPheLeuGlyLeuGlyValIleValAspAsnGlyAsnGlyAlaArgValGln 62			
Db	159 CCTGGGCTGCTCAGGAGCAGAGT-----GGCACTGGCTCAAAACAGGG 203			
Qy	63 ArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThr 82			
Db	204 TAAGCTCGGGGCGCTACCTCTCCAGCAGTCTCGA-----ACAGGCTCTCTTGTC 257			
Qy	83 AlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHis 102			

Best Local Similarity:	27.14%	Mismatches:	52
Query Match:	14.44%	Indels:	33
DB:	12	Gaps:	6

US-09-684-215A-4 (1-132) x BI350520 (1-499)

QY	9	LeuSerGlnGly-----	-----GlyClnGlyPheAlaIleProIle	20
			::: :	
Db	430	CTGGAACAGCGCCTCTGGACACCTGAACCTTTTCATTGCCCAATGGTATGACCTGTCCGC TC	371	
QY	21	GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis	40	
		::::::		
Db	370	AGGCAGCTGTGTACTCTT-----CTCCAGGGAGGAGGAAAGCAGCGTGCCCAT	320	
QY	41	Ile-----	-----GlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsn	55
Db	319	CTCTTGCTCAAAGTCAAGGCCACATAGACAGACTTCTCTTTGATGTACCGACACA---	263	
QY	56	GlyAsnGlyAlaArgValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIle	75	
Db	262	-----TTCCCTCTCAGCTGTGTGGTGAAGCTGTACGCTCTCTCGGTACG	218	
QY	76	SerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMet	95	
		:::    :::     ::: :	:::	
Db	217	GATCTTTCATGAGGTAGTCAGTCAAGTCCAGGCCAGCCAAGTCCAGACGAGGAGTGGCATG	158	
QY	96	AlaAspAla-----	-----LeuAsnGlyHisHisProGlyAspValIleSerValThrTrp	112
		::::	:::	
Db	157	GGCGAGGCGGTAAACCCTGTAGATGGCACAGTGTGGGTGACACCATCACCATATGCCAT	98	
QY	113	GlnThrLysSerGlyThrArg-----	-----ThrGlyAsnVal	124
Db	97	CACGATACCACTGTGTAGCAGTACAGAGGCGATACAGAAACAGCACAGCCTGTGATGGCAACGTG	38	

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 5, 2003, 07:24:34 ; Search time 58.2414 Seconds  
(without alignments)  
1390.474 Million cell updates/sec

Title: us-09-684-215a-17

Perfect score: 148

Sequence: 1 TAASDNFQLSQGQGFPAIPGQAWIAGQI 30

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xip  
-Q=cgn2.1/USPTO.spool\_P/US09684215/runat\_05092003.072201.548/app\_query.fasta\_1.853  
-DB=N.Geneseq.19Jun03 -CPMT=fastap -SUFFIX=ring -MINMATCH=0 -1-LOOPCL=0  
-LOOPEXT=0 -UNITS=bits START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi  
-LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09684215 -RCGN\_1\_1\_626\_runat\_05092003.072201.548 -NCFU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N.Geneseq.19Jun03.\*  
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*  
25: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	148	100.0	186	24	ABK39777 DNA encoding Ral2S
2	148	100.0	186	25	ACAL12106 Human His-tagged R
3	148	100.0	186	25	ACAO3292 Lung cancer therap
4	148	100.0	396	23	AA40769 Nucleotide sequenc
5	148	100.0	399	24	ABSL7801 Human DNA encoding
6	148	100.0	447	18	AA91486 Mycobacterium tube
7	148	100.0	447	18	AA91403 Mycobacterium tube
8	148	100.0	447	19	AAV64450 M. tuberculosis im
9	148	100.0	447	19	AAV44342 Mycobacterium tube
10	148	100.0	447	20	AA219252 M. tuberculosis an
11	148	100.0	447	20	AA219040 M. tuberculosis re
12	148	100.0	447	22	AA03780 M. tuberculosis DN
13	148	100.0	447	24	AA047080 Mycobacterium tube
14	148	100.0	447	24	AA028339 Mycobacterium spec
15	148	100.0	672	23	AA040772 Nucleotide sequenc
16	148	100.0	675	22	AA064132 Human /M. tubercul
17	148	100.0	675	22	AAH93896 Ral2-P510S-C const
18	148	100.0	675	24	ABL95503 Ral2- P510S-C const
19	148	100.0	675	25	ACA59940 Prostate cancer th
20	148	100.0	702	20	AA220206 Mycobacterium tube
21	148	100.0	702	23	AA040770 Nucleotide sequenc
22	148	100.0	702	24	ABK14140 DNA encoding anti
23	148	100.0	822	24	ABK39769 DNA encoding lung
24	148	100.0	822	25	ACAL2098 Human lung cancer
25	148	100.0	822	25	ACA03284 Lung cancer therap
26	148	100.0	861	24	ABK39776 DNA encoding Ral2S
27	148	100.0	861	25	ACAL12105 Human Ral2S-L985P
28	148	100.0	861	25	ACA03291 Lung cancer therap
29	148	100.0	894	24	ABK27798 Human colon cancer
30	148	100.0	900	24	ABQ92443 Human lung cancer
31	148	100.0	915	22	AA041414 Ral2/C-terminal po
32	148	100.0	915	22	AA041414 Human /M. tubercul
33	148	100.0	915	22	AAH93905 Ral2-P775P-ORF3 co
34	148	100.0	915	24	ABL95512 Ral2- P775P-ORF3 co
35	148	100.0	915	25	ACA59949 Prostate cancer th
36	148	100.0	945	24	ABK39768 DNA encoding lung
37	148	100.0	945	25	ACAL2097 Human lung cancer
38	148	100.0	945	25	ACA03283 Lung cancer therap
39	148	100.0	1002	24	AA047077 Mycobacterium tube
40	148	100.0	1002	24	AA047078 Mycobacterium tube
41	148	100.0	1002	24	AA028336 Mycobacterium sp.
42	148	100.0	1002	24	AA028337 Mycobacterium spec
43	148	100.0	1012	24	ABQ92442 Human lung cancer
44	148	100.0	1012	24	ABL49256 Ral2/N-terminal po
45	148	100.0	1035	24	ABK69714 DNA encoding human

ALIGNMENTS

RESULT 1

ABK39777

ID ABK39777 standard; cDNA; 186 Bp.

XX ABK39777;

AC ABK39777;

XX 21-MAY-2002 (first entry)

DT DNA encoding Ral2S-L985PEx peptide.

DE Lung tumour; cancer; T cell; immune response stimulator;

XX cytostatic; gene; SS.

OS Homo sapiens.

XX Synthetic.

XX WO200204514-A2.

PN

XX 17-JAN-2002.  
 XX 10-JUL-2001; 2001WO-US22058.  
 XX 11-JUL-2000; 2000US-0614124.  
 PR 29-AUG-2000; 2000US-0651563.  
 PR 08-SEP-2000; 2000US-0658824.  
 PR 26-SEP-2000; 2000US-0671325.  
 PR 06-OCT-2000; 2000US-0677419.  
 PR 30-OCT-2000; 2000US-0702705.  
 PR 13-DEC-2000; 2000US-0736457.  
 PR 03-MAY-2001; 2001US-0849626.  
 XX (CORI-) CORIXA CORP.  
 XX Wang T, Watanabe Y, Henderson RA, Johnson JC, Retter MW;  
 PI Marnerakis M, Carter D, Fanger GR, Vedvick IS, Bangur CS;  
 PI McNabb A, Wang A, Fanger N, Switzer A, McNeill PD, Clapper JD;  
 XX WPI; 2002-164634/21.  
 DR P-PSDB; AAU85595.  
 XX Novel polynucleotide encoding a lung tumour polypeptide useful for  
 PT stimulating and/or expanding T cells specific for a tumour protein  
 XX Claim 1; SEQ ID No 1879; 223pp; English.  
 XX The invention describes an isolated polynucleotide and polypeptide  
 CC useful for stimulating and/or expanding T cells specific for a tumour  
 CC protein for determining the presence of a cancer in a patient. A  
 CC composition containing the polynucleotide and/or polypeptide is useful  
 CC for treating a lung cancer in a patient. The polypeptide is useful for  
 CC removing tumour cells from a biological sample. The polynucleotide is  
 CC also useful as probe or primer to detect the level of mRNA encoding a  
 CC tumour protein. This sequence encodes a lung tumour associated protein  
 CC or protein fragment, described in the method of the invention.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX SQ Sequence 186 BP; 39 A; 61 C; 47 G; 39 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 2,99e-15 Length: 186  
 Score: 148.00 Matches: 30  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 24 Gaps: 0  
 US-09-684-215A-17 (1-30) x ABK39777 (1-186)  
 QY 1 ThrAlaAlaSerAspAspPheClnLeuSerGlnGlnGlyGlnGlyPheAlaIleProIle 20  
 DB 22 AGCGCGCGTCCGATPACTCCAGCTGCCAGGTGGCAGGATTCGCCATTCGCATC 81  
 QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30  
 DB 82 GGCAGCGCGATGGCGATCGCGGCCAGATC 111  
 RESULT 2  
 ID ACA12106  
 XX ACA12106 standard; cDNA; 186 BP.  
 AC ACA12106;  
 XX ACA12106;  
 DT 06-JUN-2003 (first entry)  
 XX Human His-tagged Ral2S-L985pEx fusion protein cDNA.  
 DE Human; lung cancer; ss; lung tumour; cytostatic; vaccine;  
 KW

KW T cell expansion; CD4; CD8; RAL2; gene.  
 XX Homo sapiens.  
 OS Synthetic.  
 XX US2002197669-A1.  
 PN 26-DEC-2002.  
 PD 03-MAY-2001; 2001US-0849626.  
 PF 13-DEC-2000; 2000US-0736457.  
 PR (BANG/) BANGUR C S.  
 PA (FANG/) FANGER G R.  
 PA (WANG/) WANG A.  
 PA (WANG/) WANG T.  
 PA (SWIT/) SWITZER A P.  
 PA (MCNE/) MCNEILL P D.  
 PA (CLAP/) CLAPPER J D.  
 XX Bangur CS, Fanger GR, Wang A, Wang T, Switzer AP, McNeill PD;  
 PI Clapper JD;  
 PI WPI; 2003-352750/33.  
 DR P-PSDB; ABU69570.  
 XX Novel lung cancer polynucleotide encoding lung cancer protein, useful  
 PT for detecting the presence of lung cancer in a patient, and in  
 PT pharmaceutical compositions, e.g. vaccines, for treating lung cancer  
 XX Example 10; Page -: 72pp; English.  
 XX The invention relates to a polynucleotide encoding a lung tumour protein,  
 CC comprising a sequence selected from any of the 14 sequences  
 CC mentioned in the specification, or a sequence (S2) mentioned in  
 CC specification, complement of S1, sequences consisting of at least 20  
 CC contiguous residues of S1, sequences that hybridise to S1, sequences  
 CC having 75%, preferably 90%, identity to S1, or degenerate variants of  
 CC S1. Also included are an isolated polypeptide (comprising a sequence (S3)  
 CC selected from any one of the 4 amino acid sequences mentioned in the  
 CC specification, a sequence encoded by the polynucleotide, or sequences  
 CC having at least 70%, preferably 90%, identity to a sequence encoded by  
 CC the polynucleotide), an expression vector comprising the polynucleotide  
 CC operably linked to an expression control sequence, a host cell  
 CC transformed or transfected with the vector, an isolated antibody (or its  
 CC antigen-binding fragment) that specifically binds to the polypeptide,  
 CC detecting the presence of a cancer in a patient, a fusion protein  
 CC comprising the polypeptide, an oligonucleotide that hybridises to  
 CC S1 under moderately stringent conditions, stimulating and/or expanding T  
 CC cells specific for a tumour protein (comprising contacting T cells with  
 CC the polynucleotide, protein or antigen-presenting cells, under conditions  
 CC and for a time sufficient to permit the stimulation and/or expansion of T  
 CC cells) and inhibiting the development of a cancer in a patient (by  
 CC incubating CD4<sup>+</sup> and/or CD8<sup>+</sup> T cells isolated from a patient with the  
 CC polynucleotide, protein or antigen presenting cells that express the  
 CC polynucleotide, such that T cells proliferate, administering to the  
 CC patient an effective amount of the proliferated T cells, and thus  
 CC inhibiting the development of a cancer in the patient. The  
 CC polynucleotide, protein and cells are useful in a composition for  
 CC stimulating an immune response in a patient, and for treating a cancer in  
 CC a patient (particularly lung cancer). The oligonucleotide is useful for  
 CC determining the presence of a cancer in a patient. The protein and  
 CC oligonucleotides are useful in pharmaceutical compositions, e.g.  
 CC vaccines. The polynucleotide is also useful as a probe or primer for  
 CC nucleic acid hybridisation, and in the design and preparation of  
 CC ribozyme molecules for inhibiting expression of tumour polypeptides and  
 CC proteins in tumour cells. An amplified portion of the polynucleotide is  
 CC useful for isolating a full-length gene from a suitable library.  
 CC The present sequence encodes a fusion protein of human RAL2 with  
 CC the protein product of a cDNA (full length, extended or partial)  
 CC isolated from a library derived from lung tumour/cancer cells.  
 CC Note: The sequence data for this patent did not form part

CC of the printed specification, but was obtained in electronic  
 CC format directly from the USPTO  
 CC at seqdata.uspto.gov/sequence.html?DocId=20020197669.

XX Sequence 186 BP; 39 A; 61 C; 47 G; 39 T; 0 other;

Alignment Scores: 2.99e-15 Length: 186  
 Pred. No.: 148.00 Matches: 30  
 Score: 148.00  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 25 Gaps: 0

US-09-684-215A-17 (1-30) x ACA12106 (1-186)

Qy 1 ThrAlaAlaSerAspAsnGlnLeuSerGlnGlyGlnGlyPheAlaIleProile 20  
 Db 22 ACGCCGCGTCCGATACCTTCCAGCTGTCCAGGGTGGCAGGATTCGCCATTCCGATC 81  
 Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30  
 Db 82 GGGCAGCGGATGGCGATCGCGGGCCAGATC 111

RESULT 3

ACA03292  
 ID ACA03292 standard; cDNA: 186 BP.

XX ACA03292;

XX 22-MAY-2003 (first entry)

XX Lung cancer therapy and diagnosis associated cDNA #1767.

XX Lung cancer; cytostatic; vaccine; gene therapy; cancer;

XX gene; ss.

XX Homo sapiens.

XX US2002172952-A1.

XX 21-NOV-2002.

XX 10-JUL-2001; 2001US-0902941.

XX 30-JUN-1999; 99US-0346492.

XX 15-OCT-1999; 99US-0419356.

XX 17-DEC-1999; 99US-0466867.

XX 30-DEC-1999; 99US-0476300.

XX 06-MAR-2000; 2000US-0519642.

XX 22-MAR-2000; 2000US-0533077.

XX 10-APR-2000; 2000US-0546259.

XX 27-APR-2000; 2000US-0560406.

XX 05-JUN-2000; 2000US-0589184.

XX 11-JUL-2000; 2000US-0614124.

XX 29-AUG-2000; 2000US-0651563.

XX 08-SEP-2000; 2000US-0658824.

XX 26-SEP-2000; 2000US-0671325.

XX 06-OCT-2000; 2000US-0677419.

XX 30-OCT-2000; 2000US-0702705.

XX 13-DEC-2000; 2000US-0736457.

XX 03-MAY-2001; 2001US-0849626.

XX (CORI-) CORIXA CORP.

PS Example 10; SEQID NO 1879; 82pp; English.

XX The invention describes an isolated polynucleotide comprising one of 32  
 CC sequences, complement or degenerate variants of them. The polynucleotide  
 CC is useful for preparing a composition e.g. a vaccine or for gene therapy,  
 CC for treating or inhibiting development of cancer, e.g. lung cancer.  
 CC This sequence represents a polynucleotide associated with the  
 CC compositions and methods for the therapy and diagnosis of lung cancer.

XX Sequence 186 BP; 39 A; 61 C; 47 G; 39 T; 0 other;

Alignment Scores: 2.99e-15 Length: 186  
 Pred. No.: 148.00 Matches: 30  
 Score: 148.00  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 25 Gaps: 0

US-09-684-215A-17 (1-30) x ACA03292 (1-186)

Qy 1 ThrAlaAlaSerAspAsnGlnLeuSerGlnGlyGlnGlyPheAlaIleProile 20  
 Db 22 ACGCCGCGTCCGATACCTTCCAGCTGTCCAGGGTGGCAGGATTCGCCATTCCGATC 81  
 Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30  
 Db 82 GGGCAGCGGATGGCGATCGCGGGCCAGATC 111

RESULT 4

AAL40769  
 ID AAL40769 standard; DNA; 396 BP.

XX AAL40769;

XX 03-OCT-2002 (first entry)

XX Nucleotide sequence encoding Ral2 protein.

XX Ral2; serine protease antigen; MTB32A; Mycobacterium tuberculosis;  
 KW vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;  
 KW immunogen; cytokine; gene; ds.

XX Unidentified.

XX Key Location/Qualifiers

XX CDS 1..396

XX FT /\*tag= a

XX FT /product= "Ral2 protein"

XX PN /note= "No start or stop codon"

XX WO200125401-A2.

XX 12-APR-2001.

XX 06-OCT-2000; 2000WO-US27652.

XX 07-OCT-1999; 99US-0158585.

XX (CORI-) CORIXA CORP.

XX Skeiky Y, Guderian J;

XX WPI; 2001-266299/27.

XX P-PSDB; RAO22138.

XX Recombinant nucleic acid molecule for producing high yield expression  
 PT of desired fusion polypeptides, encodes fusion polypeptide comprising  
 PT Mycobacterium tuberculosis coding sequence and heterologous polypeptide

XX Claim 1; Fig 2; 39pp; English.

XX

CC The invention relates to a recombinant nucleic acid molecule encoding a  
 CC fusion polypeptide, comprising a polynucleotide sequence of Ral2, a 14  
 CC kDa C-terminal fragment of serine protease antigen MTB32A of  
 CC Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.  
 CC The recombinant fusion nucleic acids and polypeptides are useful for  
 CC providing stable and high yield expression of fusion polypeptides of both  
 CC eukaryotic and prokaryotic origin and to encode a protein product for use  
 CC as an antigen for detecting serum antibodies. The presence of serum  
 CC antibodies to M. tuberculosis antigens in an individual indicates that  
 CC the individual is infected with it. The fusion polypeptides are useful as  
 CC sources of proteins for monitoring binding of serum antibodies to fusion  
 CC proteins and as an immunogen to induce and/or enhance immune responses.  
 CC The coding sequences can be ligated with a coding sequence of another  
 CC molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and  
 CC can be used in vivo as a DNA vaccine. This polynucleotide sequence  
 CC represents the DNA encoding the Ral2 protein.

XX Sequence 396 BP; 68 A; 132 C; 134 G; 62 T; 0 other;

#### Alignment Scores:

Pred. No.:	7.59e-15	Length:	396
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	23	Indels:	0
DB:		Gaps:	0

US-09-684-215A-17 (1-30) x AAL40769 (1-396)

QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20

Db 1 ACGGCCGGCTCCGATAAAGTCCAGCTGCCAGGCTGGCAGGGATTGCCATTCGGATC 60

QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30

Db 61 GGGCAGGCGATGGCGATCGCGGGCCAGATC 90

#### RESULT 5

ABS71801

ID ABS71801 standard; DNA; 399 BP.

XX ABS71801;

DT 02-DEC-2002 (first entry)

XX Human DNA encoding a mammaglobin/Ral2 fusion protein.

XX Human; epitope: mammaglobin; breast cancer; cytostatic; T cell; CD4+;  
 KW CD8+; antigen; Ral2; ds; gene.

OS Chimeric - Homo sapiens.

OS Chimeric - Mycobacterium tuberculosis.

PN WO200253017-A2.

XX 11-JUL-2002.

PF 08-JAN-2002; 2002WO-US03057.

PR 08-JAN-2001; 2001US-0757417.

PR 08-NOV-2001; 2001US-0008045.

XX (CORI-) CORIXA CORP.

XX Fling SP, Foy TM, Clapper JD, Wang A, Johnson JC, McNeill PD;  
 PI Sutherland RA;

DR WPI; 2002-706844/76.

DR P-PSDB; ABG94684.

XX Novel polypeptides comprise one or more human mammaglobin epitopes and  
 PT polynucleotides encoding the polypeptides, useful for preventing and  
 PT treating breast cancers -

XX Claim 20; Page 115; 121pp; English.  
 PS The invention relates to an isolated polypeptide (I) comprising 7-30  
 CC consecutive amino acid residues of human mammaglobin, where one or  
 CC more mammaglobin-specific T cells specifically reacts with (I).  
 CC Also included are (1) a composition comprising (I), in combination with  
 CC a physiologically acceptable carrier or immunostimulant; (2) a diagnostic  
 CC kit, comprising (I) and a detection reagent comprising a reporter group;  
 CC (3) removing tumour cells from a biological sample, by contacting a  
 CC biological sample with T cells that specifically react with (I), under  
 CC conditions and for a time sufficient to permit the removal of cells  
 CC expressing mammaglobin or a peptide epitope from the sample; (4) an  
 CC isolated T cell population (II), comprising T cells prepared using (I);  
 CC and (5) a composition (C2) comprising a polynucleotide comprising a  
 CC sequence encoding human mammaglobin (or variant or tagged with an  
 CC affinity tag), or a polypeptide comprising a human mammaglobin (or  
 CC variant or tagged with an affinity tag), in combination with an  
 CC immunostimulant. The peptide is useful for inhibiting the development of  
 CC breast cancer in a patient, by incubating CD4+ and/or CD8+ T cells  
 CC isolated from a patient with a mammaglobin epitope, such that T cells  
 CC proliferate, administering the proliferated T cells to the patient,  
 CC optionally cloning at least one proliferated T cell and administering at  
 CC least one cloned cell, and thus inhibiting the development of breast  
 CC cancer in the patient. The compositions and methods are useful for  
 CC inhibiting the development of breast cancer in a patient. The  
 CC present sequence encodes a human mammaglobin/Ral2 fusion protein.

XX Sequence 399 BP; 109 A; 96 C; 92 G; 102 T; 0 other;

#### Alignment Scores:

Pred. No.:	7.66e-15	Length:	399
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	24	Indels:	0
DB:		Gaps:	0

US-09-684-215A-17 (1-30) x ABS71801 (1-399)

QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20

Db 22 ACGGCCGGCTCCGATAAAGTCCAGCTGCCAGGCTGGCAGGGATTGCCATTCGGATC 81

QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30

Db 82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

#### RESULT 6

AAT91466

ID AAT91466 standard; DNA; 447 BP.

XX AAT91466;

DT 07-JAN-1998 (first entry)

XX Mycobacterium tuberculosis antigen TBra12 encoding DNA.

XX Antigen; immunogen; vaccine; tuberculosis; non specific adjuvant;  
 KW skin testing; M. tuberculosis; ss.

OS Mycobacterium tuberculosis.

XX Key Location/Qualifiers

FT CDS 11..409

FT /\*tag= a

FT /product= Antigen\_TBra12

XX WO9709428-A2.

XX 13-MAR-1997.

XX 30-AUG-1996; 96WO-US14674.

```
XX PR 12-JUL-1996; 96US-0680574.
XX PR 01-SEP-1995; 95US-0523436.
XX PR 22-SEP-1995; 95US-0533634.
XX PR 22-MAR-1996; 96US-0620874.
XX PR 05-JUN-1996; 96US-0659683.
XX PA (CORI-) CORIXA CORP.
XX XX Campos-neto A, Dillon DC, Houghton R, Reed SG, 'Skeiky YAW;
XX PI Twardzik DR, Vedvick TH;
XX DR WPI; 1997-192903/17.
XX DR P-PSDB; AAW32422.
XX XX New immunogenic polypeptide(s) from Mycobacterium tuberculosis - are
XX PT useful in vaccines for prevention or treatment of tuberculosis, also
XX PT for diagnosis
XX PS Claim 3; Page 50; 168pp; English.
XX XX A new immunogenic polypeptide has been developed comprising an
XX CC immunogenic part of a soluble Mycobacterium tuberculosis antigen (or
XX CC its variant differing only in conservative substitutions and/or
XX CC modifications). The present sequence encodes a specifically claimed
XX CC M.tuberculosis antigen, TbrA12. The immunogenic protein, and fusion
XX CC proteins containing one or more of the proteins or one of the proteins
XX CC plus ESAT-6, are useful in vaccines, preferably when formulated with a
XX CC non-specific adjuvant, to induce an immune response against
XX CC M.tuberculosis (for treatment or prevention).
XX XX Sequence 447 BP; 79 A; 146 C; 149 G; 72 T; 1 other;
SQ Alignment Scores:
Pred. No.: 8.81e-15 Length: 447
Score: 148.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0
US-09-684-215A-17 (1-30) x AAT91466 (1-447)
Qy 1 ThrAlaIaSerAspAsnPhGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 11 ACGGCGCGTCCGATACCTCCAGCTGTCCAGGGTGGCGAGGATTCCGCATTCCGATC 70
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
Db 71 GGGCAGCGGATGGCGATCGCGGCCAGATC 100
RESULT 7
AAT91403
ID AAT91403 standard; DNA; 447 BP.
XX AC AAT91403;
XX DT 12-JAN-1998 (first entry)
XX XX Mycobacterium tuberculosis antigen TbrA12 encoding DNA.
XX DE Antigen; immunogen; vaccine; tuberculosis; non specific adjuvant;
XX KW skin testing; M.tuberculosis; ss.
XX OS Mycobacterium tuberculosis.
XX FH Key Location/Qualifiers
XX FT CDS 11..409
XX FT /*tag= a
XX FT /product= Antigen_TbrA12
XX PN W09709429-A2.
XX XX
```

```
PD 13-MAR-1997.
XX 30-AUG-1996; 96WO-US14675.
XX 12-JUL-1996; 96US-0680573.
XX PR 01-SEP-1995; 95US-0523435.
XX PR 22-SEP-1995; 95US-0532136.
XX PR 22-MAR-1996; 96US-0620280.
XX PR 05-JUN-1996; 96US-0658800.
XX XX (CORI-) CORIXA CORP.
XX XX Campos-neto A, Dillon DC, Houghton R, Reed SG, Skeiky YAW;
XX PI Twardzik DR, Vedvick TH;
XX XX WPI; 1997-192904/17.
XX DR P-PSDB; AAW32354.
XX XX New immunogenic polypeptide(s) from soluble M. tuberculosis antigens
XX PT - useful for diagnosis of M. tuberculosis infection
XX PS Claim 3; Page 49-50; 190pp; English.
XX XX A new immunogenic polypeptide has been developed comprising an
XX CC immunogenic part of a soluble Mycobacterium tuberculosis antigen (or
XX CC its variant differing only in conservative substitutions and/or
XX CC modifications). The present sequence encodes a specifically claimed
XX CC M.tuberculosis antigen, TbrA12. The immunogenic polypeptide can be
XX CC used to diagnose M.tuberculosis infection by forming complexes with
XX CC specific antibodies in the sample. Fragments of DNA encoding the
XX CC immunogenic polypeptide can be used as diagnostic primers or probes
XX CC and agents that bind to the antigen, especially monoclonal antibodies
XX CC or equivalent polyclonal antibodies, are also used for diagnosis.
XX SQ Sequence 447 BP; 79 A; 146 C; 149 G; 72 T; 1 other;
Alignment Scores:
Pred. No.: 8.81e-15 Length: 447
Score: 148.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0
US-09-684-215A-17 (1-30) x AAT91403 (1-447)
Qy 1 ThrAlaIaSerAspAsnPhGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 11 ACGGCGCGTCCGATACCTCCAGCTGTCCAGGGTGGCGAGGATTCCGCATTCCGATC 70
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
Db 71 GGGCAGCGGATGGCGATCGCGGCCAGATC 100
RESULT 8
AAV64450
ID AAV64450 standard; DNA; 447 BP.
XX AC AAV64450;
XX DT 27-JAN-1999 (first entry)
XX XX M. tuberculosis immunogenic polypeptide TbrA12 DNA.
XX DE Tuberculosis; immunogenic; soluble; antigen; protective immunity; TB;
XX KW vaccine; pharmaceutical; infection; diagnosis; ss.
XX OS Mycobacterium tuberculosis.
XX PN W09816646-A2.
XX XX
XX PD 23-APR-1998.
XX XX
```

```
PF 07-OCT-1997; 97WO-US18293.
XX
PR 13-MAR-1997; 97US-0818112.
PR 11-OCT-1996; 96US-0730510.
XX
PA (CORI-) CORIXA CORP.
XX
PI Campos-Neto A, Dillon DC, Houghton R, Lodes MJ;
PI Reed SG, Skeiky YAW, Twardzik DR, Vedvick TS;
DR WPI; 1998-261042/23.
DR P-PSDB; AAW81657.
XX
PT Immunogenic Mycobacterium tuberculosis polypeptide(s) and DNA - used
PT to develop products for the detection of M. tuberculosis infection
PT and for diagnosis, treatment and prevention of tuberculosis
XX
PS Claim 3; Page 62-63; 230pp; English.
XX
CC This sequence encodes an immunogenic portion of a soluble Mycobacterium
CC tuberculosis (MT) antigen which can be used in a method for inducing
CC protective immunity against tuberculosis (TB). This sequence can be
CC formulated into vaccines and/or pharmaceutical compositions for
CC immunising against M. tuberculosis infection or may be used for the
CC diagnosis of tuberculosis.
XX
SQ Sequence 447 BP; 79 A; 146 C; 149 G; 72 T; 1 other;

Alignment Scores:
Pred. No.: 8,81e-15 Length: 447
Score: 148.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 19 Indels: 0
DB: Gaps: 0

US-09-684-215A-17 (1-30) x AAV64450 (1-447)
QY 1 ThrAlaLaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20
DB 11 ACGGCCGCTCGATAACTTCCAGCTGTCCAGGGTGGCAGGGATTCCGCATTCGGATC 70
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
DB 71 GGGCAGGCGATGGCGATCGCGGCCAGATC 100

RESULT 9
ID AAV44342
XX AAV44342 standard; DNA; 447 BP.
XX
AC AAV44342;
XX
DT 09-NOV-1998 (first entry)
XX
DE Mycobacterium tuberculosis antigen TBRa12 DNA.
XX
KW Tuberculosis; infection; diagnosis; antigen; TBRa12; ss.
XX
OS Mycobacterium tuberculosis strain H37Ra.
FH Location/Qualifiers
FT 11..406
FT /*tag= a
XX
PN WO9816645-A2.
XX
PD 23-APR-1998.
XX
PF 07-OCT-1997; 97WO-US18214.
XX
PR 13-MAR-1997; 97US-0818111.
PR 11-OCT-1996; 96US-0729622.
XX

(PORI-) CORIXA CORP.
XX
PI Campos-Neto A, Dillon DC, Houghton R, Lodes MJ;
PI Reed SG, Skeiky YAW, Twardzik DR, Vedvick TS;
DR WPI; 1998-251292/22.
DR P-PSDB; AAW64294.
XX
PT New isolated Mycobacterium tuberculosis polypeptides and DNA - used
PT to develop products for the detection of M. tuberculosis infection
PT and diagnosis of tuberculosis
XX
PS Claim 3; Page 61; 250pp; English.
XX
CC This DNA sequence codes for Mycobacterium tuberculosis soluble
CC antigen TBRa12 (see AAW64294). It was isolated from a M. tuberculosis
CC strain H37Ra expression library with rabbit anti-sera raised
CC against M. tuberculosis supernatant. The invention relates to
CC compositions and methods for diagnosing tuberculosis. It provides
CC polypeptides (see AAW64291-W64379) comprising an antigenic portion of
CC a soluble M. tuberculosis antigen, or an immunogenic portion of an
CC M. tuberculosis antigen, as well as DNA sequences encoding such
CC polypeptides, recombinant expression vectors and transformed or
CC transfected host cells. Also claimed are methods and diagnostic
CC kits for detecting M. tuberculosis infection in a patient using
CC these polypeptides, antibodies or oligonucleotide probes and
CC primers, for the diagnosis of tuberculosis.
XX
SQ Sequence 447 BP; 79 A; 146 C; 149 G; 72 T; 1 other;

Alignment Scores:
Pred. No.: 8,81e-15 Length: 447
Score: 148.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 19 Indels: 0
DB: Gaps: 0

US-09-684-215A-17 (1-30) x AAV44342 (1-447)
QY 1 ThrAlaLaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20
DB 11 ACGGCCGCTCGATAACTTCCAGCTGTCCAGGGTGGCAGGGATTCCGCATTCGGATC 70
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
DB 71 GGGCAGGCGATGGCGATCGCGGCCAGATC 100

RESULT 10
ID AA219252
XX AA219252 standard; DNA; 447 BP.
XX
AC AA219252;
XX
DT 05-NOV-1999 (first entry)
XX
DE M. tuberculosis antigen TBRa12 DNA sequence.
XX
KW Mycobacterium tuberculosis; M. tuberculosis; antigen; immunogen;
KW immunotherapy; diagnosis; immunisation; vaccine; infection;
KW immune response; skin test; ss.
XX
OS Mycobacterium tuberculosis.
XX
PN WO9942076-A2.
XX
PD 26-AUG-1999.
XX
PF 17-FEB-1999; 99WO-US03268.
XX
PR 05-MAY-1998; 98US-0072967.
PR 18-FEB-1998; 98US-0025197.
XX
```



PA (CORI-) CORIXA CORP.  
XX Campos-Neto A, Dillon DC, Hendrickson RC, Houghton R;  
PI Lodes MJ, Reed SG, Skeiky YAW, Twardzik DR, Vedvick TS;  
XX  
XX WPI; 1999-527409/44.  
DR P-PSDB; AAY39096.  
XX  
XX New antigens from Mycobacterium tuberculosis useful in diagnostic  
PT skin tests and protective or therapeutic vaccines or compositions  
XX  
XX Claim 3; Page 73; 299pp; English.  
XX  
XX The present invention describes polypeptides comprising an immunogenic  
CC part of a Mycobacterium tuberculosis antigen (Ag). Also described  
CC are vaccines and fusion protein containing M. tuberculosis Ag's.  
CC M. tuberculosis Ag's, DNAs encoding them, derived fusion proteins and  
CC other polypeptides fragments, can be used in pharmaceutical compositions  
CC or vaccines to generate a protective or therapeutic immune response to  
CC M. tuberculosis and as reagents in skin tests for diagnosis of  
CC tuberculosis. Ag can induce proliferation of, or cytokine secretion  
CC by, T, B or natural killer cells and/or macrophages in  
CC tuberculosis-immune subjects. AA219249 to AA219460 and AAY39083 to  
CC AAY39225 are used in the exemplification of the present invention.  
XX  
SQ Sequence 447 BP; 79 A; 146 C; 149 G; 72 T; 1 other;  
  
Alignment Scores:  
Pred. No.: 8,81e-15 Length: 447  
Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 20 Gaps: 0  
  
US-09-684-215A-17 (1-30) x AA219252 (1-447)  
  
QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProile 20  
DB 11 ACGCCGCGCTCCGATAACTTCCAGCTGCCAGGGTGGCGAGGATTCGCCATTCGGATC 70  
  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30  
DB 71 GGCAGCGCATGCGATCGCGGCCAGATC 100  
  
RESULT 11  
AA219040  
ID AA219040 standard; DNA: 447 BP.  
XX  
AC AA219040;  
XX  
XX 05-NOV-1999 (first entry)  
XX  
XX M. tuberculosis recombinant antigen DNA encoding TbrA12.  
XX  
XX Antigen; diagnosis; detection; infection; antibody; immunisation;  
XX vaccine; immunity; ss.  
XX  
XX Mycobacterium tuberculosis.  
XX  
XX WO9942118-A2.  
XX  
XX 26-AUG-1999.  
XX  
XX 17-FEB-1999; 99WO-US03265.  
XX  
XX 05-MAY-1998; 98US-0072596.  
XX  
XX 18-FEB-1998; 98US-0024753.  
XX  
XX (CORI-) CORIXA CORP.  
XX  
XX Campos-Neto A, Dillon DC, Hendrickson RC, Houghton R;  
PI Lodes MJ, Reed SG, Skeiky YAW, Twardzik DR, Vedvick TS;

XX  
DR WPI; 1999-527416/44.  
DR P-PSDB; AAY38959.  
XX  
XX New polypeptide comprising antigenic portions of M. tuberculosis  
XX  
XX Claim 10; Page 103; 323pp; English.  
XX  
XX This invention describes novel recombinant antigens and their encoding  
CC nucleic acids derived from Mycobacterium tuberculosis. The novel  
CC polypeptides are useful for detecting M. tuberculosis infection in a  
CC biological sample by detecting antibodies which bind with the  
CC polypeptides, and are useful as vaccines for immunizing against  
CC M. tuberculosis infection. The new detection methods are needed as  
CC current vaccination strategies do not provide 100% immunity.  
XX  
SQ Sequence 447 BP; 79 A; 146 C; 149 G; 72 T; 1 other;  
  
Alignment Scores:  
Pred. No.: 8,81e-15 Length: 447  
Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 20 Gaps: 0  
  
US-09-684-215A-17 (1-30) x AA219040 (1-447)  
  
QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProile 20  
DB 11 ACGCCGCGCTCCGATAACTTCCAGCTGCCAGGGTGGCGAGGATTCGCCATTCGGATC 70  
  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30  
DB 71 GGCAGCGCATGCGATCGCGGCCAGATC 100  
  
RESULT 12  
AA203780  
ID AA203780 standard; DNA: 447 BP.  
XX  
AC AA203780;  
XX  
XX 29-AUG-2001 (first entry)  
XX  
XX M. tuberculosis DNA encoding a partial antigen TbrA12.  
XX  
XX TbrA12; antigen; vaccine; tuberculosis; AIDS;  
XX acquired immunodeficiency disease; ss.  
XX  
XX Mycobacterium tuberculosis.  
XX  
XX Key Location/Qualifiers  
XX CDS 11..409  
XX /\*tag= a  
XX /product= "TbrA12"  
XX /partial  
XX /note= "No start codon"  
XX  
XX WO200124820-A1.  
XX  
XX 12-APR-2001.  
XX  
XX 10-OCT-2000; 2000WO-US28095.  
XX  
XX 07-OCT-1999; 99US-0158338.  
XX  
XX 07-OCT-1999; 99US-0158425.  
XX  
XX (CORI-) CORIXA CORP.  
XX  
XX Skeiky Y, Reed S, Houghton RL, McNeill PD, Dillon DC, Lodes ML;  
XX  
XX WPI; 2001-250576/30.  
XX  
XX P-PSDB; AAU01889.

XX Vaccinating against Mycobacterium tuberculosis in mammals using fusion  
PT proteins comprising combinations of heterologous antigens -  
PS  
XX Example 2; Page 162; 168pp; English.

CC The sequence encodes Mycobacterium tuberculosis TbrA12,  
CC an M. tuberculosis antigen. Compositions comprising at least  
CC 2 heterologous antigens, as a fusion protein, and vectors expressing the  
CC fusion proteins are used as vaccines to prophylactically immunise  
CC mammals (especially humans) against infection by Mycobacterium. The  
CC compositions contain at least 2 heterologous antigens that increase the  
CC serological sensitivity of individuals infected with tuberculosis, a  
CC disease frequently affecting patients with acquired immunodeficiency  
CC disease, AIDS.

XX Sequence 447 BP; 79 A; 146 C; 149 G; 72 T; 1 other;

Alignment Scores:  
Pred. No.: 8,81e-15 Length: 447  
Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 22 Gaps: 0

US-09-684-215A-17 (1-30) x AAS03780 (1-447)

OY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20  
DB 11 ACGGCCGGCTCGGTAACCTCCAGCTGTCCCGAGGTGGCGAGGATCCGCCATTCGGATC 70

OY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30  
DB 71 GGGCAGGCGATGCGATCGCGGCCAGATC 100

RESULT 13  
AAD47080  
ID AAD47080 standard; DNA; 447 BP.  
XX  
AC AAD47080;  
XX  
DT 27-JAN-2003 (first entry)  
XX  
DE Mycobacterium tuberculosis Ra12 antigen encoding DNA.  
XX  
KW Vaccine; Immunity; diagnostic agent; gene therapy; Ra12 antigen;  
KW gene; ds.  
XX  
OS Mycobacterium tuberculosis.  
XX  
FH Key Location/Qualifiers  
FT CDS 11..409  
FT /tag= a  
FT /product= "Ra12 antigenic protein"  
FT /note= "No start codon"  
FT /partial  
XX  
PN WO20027292-A2.  
XX  
PD 19-SEP-2002.  
XX  
PF 13-MAR-2002; 2002WO-US08223.  
XX  
PR 13-MAR-2001; 2001US-275837P.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
PI Skeiky Y, Brannon M, Guderian J;  
XX WPI; 2002-759844/82.  
DR P-PSDB; AAE29705.  
XX

PT New recombinant nucleic acid molecule comprising a Leishmania TSA,  
PT LeIF, M15 or 6H polynucleotide, useful as vaccine to elicit protective  
PT immunity against pathogenic microorganisms e.g. Leishmania and  
PT Mycobacterium tuberculosis  
XX  
PS Disclosure: Page 83; 155pp; English.

CC The invention relates to a recombinant nucleic acid molecule encoding a  
CC fusion polypeptide. The recombinant nucleic acid comprises a heterologous  
CC polynucleotide sequence encoding an antigen or an antigenic fragment from  
CC Mycobacterium sp. and a Leishmania polynucleotide sequence encoding a  
CC polypeptide or its fragment. The Leishmania polynucleotide is selected  
CC from TSA, LeIF, M15, and 6H polynucleotides. Sequences of the invention  
CC are used in methods for eliciting immune response in mammals. They are  
CC useful as vaccines to elicit protective immunity against pathogenic  
CC microorganisms such as Leishmania and Mycobacterium tuberculosis. Fusion  
CC polypeptides are used for enhancing the expression of polynucleotides,  
CC as in vivo diagnostic agents and for raising antibodies in a non-human  
CC animal. The invention is used in gene therapy. The present sequence is  
CC M. tuberculosis Ra12 (C-terminus of MTB32A; RA35FL) antigen encoding DNA.

XX Sequence 447 BP; 79 A; 146 C; 149 G; 72 T; 1 other;

Alignment Scores:  
Pred. No.: 8,81e-15 Length: 447  
Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 22 Gaps: 0

US-09-684-215A-17 (1-30) x AAD47080 (1-447)

OY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20  
DB 11 ACGGCCGGCTCGGTAACCTCCAGCTGTCCCGAGGTGGCGAGGATTCGCCATTCGGATC 70

OY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30  
DB 71 GGGCAGGCGATGCGATCGCGGCCAGATC 100

RESULT 14  
AAD28339  
ID AAD28339 standard; cDNA; 447 BP.  
XX  
AC AAD28339;  
XX  
DT 22-APR-2002 (first entry)  
XX  
DE Mycobacterium species MTB32A C-terminal peptide encoding cDNA, MTB32A.  
XX  
KW Fusion protein; antigen; serological sensitivity; immune response;  
KW tuberculosis; infection; vaccine; MTB32A; Ra32FL; MTB32A; ss.  
XX  
OS Mycobacterium sp.  
XX  
FH Key Location/Qualifiers  
FT CDS 11..409  
FT /tag= a  
FT /product= "Ra35 protein fragment"  
FT /note= "CDS does not include start codon"  
FT /partial  
XX  
PN WO200198460-A2.  
XX  
PD 27-DEC-2001.  
XX  
PF 20-JUN-2001; 2001WO-US19959.  
XX  
PR 20-JUN-2000; 2000US-0597796.  
XX  
PR 01-FEB-2001; 2001US-265737P.  
XX  
PA (CORI-) CORIXA CORP.

XX Skeiky Y, Reed S, Alderson M;  
PI WPI: 2002-147798/19.  
DR P-PSDB; AAE17569.  
XX  
XX Composition comprising MTB39 antigen and MTB32A antigen from  
PT Mycobacterium species, useful for eliciting immune response in a  
PT subject -  
XX  
XX Disclosure; Page 99; 136pp; English.  
XX  
XX The present invention relates to fusion proteins containing at least  
CC two Mycobacterium species antigens, nucleotides encoding them and  
CC compositions comprising such fusion proteins. The present invention  
CC particularly relates to nucleic acids encoding fusion proteins that  
CC include two or more individual M. tuberculosis antigens which increase  
CC the serological sensitivity of sera from individuals infected with  
CC tuberculosis and methods for their use in diagnosis, prevention and  
CC treatment of tuberculosis infection. Sequences of the invention are  
CC useful for eliciting an immune response in a mammal, e.g., human,  
CC immunised with BCG. They are useful in the diagnosis, treatment and  
CC prevention of Mycobacterium infection. The fusion proteins and the  
CC polynucleotides are useful as diagnostic tools in patients infected  
CC with Mycobacterium, in vitro and in vivo assays for detecting humoral  
CC antibodies or cell-mediated immunity against M. tuberculosis, for the  
CC diagnosis of an infection or monitoring of disease progression, as  
CC immunogens to generate or elicit a protective immune response in a  
CC patient and for raising anti-M. tuberculosis antibodies in a non-human  
CC animal. Sequences of the invention are also used as vaccines. MTB32A  
CC fusion proteins of the invention are useful as in vivo diagnostic agents  
CC for intradermal skin test. The present sequence is Mycobacterium species  
CC MTB32A (Ra32FL) C-terminal protein fragment (residues 244-355) encoding  
CC cDNA, MTBRa12.  
XX  
XX Sequence 447 BP; 79 A; 146 C; 149 G; 72 T; 1 other;

Alignment Scores:  
Pred. No.: 8,81e-15 Length: 447  
Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 24 Gaps: 0

US-09-684-215A-17 (1-30) x AAD28339 (1-447)

QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProile 20  
DB 11 ACGGCCGCGTCCGATACCTCCAGCTGCCAGGGTGGCGAGGATTCCGCATTCGCATC 70  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30  
DB 71 GGGCAGCGATGGCGATCGCGGCCAGATC 100

RESULT 15  
AAL40772  
ID AAL40772 standard; DNA; 672 BP.  
XX  
XX AAL40772;  
XX  
XX 03-OCT-2002 (first entry)  
XX  
XX Nucleotide sequence encoding Ra12-mammaglobin fusion protein.  
DE  
XX  
XX Ra12: serine protease antigen; MTB32A; Mycobacterium tuberculosis;  
KW vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;  
KW immunogen; cytokine; gene; ds.  
XX  
XX ChimERIC - Mammalian.  
OS ChimERIC - Unidentified.  
XX  
XX  
XX Key Location/Qualifiers

FT CDS 4..666  
FT /\*tag= a  
FT /product= "Ra12-mammaglobin fusion protein"  
XX  
XX  
XX  
XX W0200125401-A2.  
XX  
XX 12-APR-2001.  
XX  
XX 06-OCT-2000; 2000WO-US27652.  
XX  
XX 07-OCT-1999; 99US-0158585.  
XX  
XX (CORI-) CORTXA CORP.  
XX  
XX Skeiky Y, Guderian J;  
XX  
XX WPI: 2001-266299/27.  
XX P-PSDB; AAO22141.  
XX  
XX Recombinant nucleic acid molecule for producing high yield expression  
PT of desired fusion polypeptides, encodes fusion polypeptide comprising  
PT Mycobacterium tuberculosis coding sequence and heterologous polypeptide  
PT  
XX  
XX Disclosure; Fig 5; 39pp; English.  
XX  
XX The invention relates to a recombinant nucleic acid molecule encoding a  
CC fusion polypeptide, comprising a polynucleotide sequence of Ra12, a 14  
CC kDa C-terminal fragment of serine protease antigen MTB32A of  
CC Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.  
CC The recombinant fusion nucleic acids and polypeptides are useful for  
CC providing stable and high yield expression of fusion polypeptides of both  
CC eukaryotic and prokaryotic origin and to encode a protein product for use  
CC as an antigen for detecting serum antibodies. The presence of serum  
CC antibodies to M. tuberculosis antigens in an individual indicates that  
CC the individual is infected with it. The fusion polypeptides are useful as  
CC sources of proteins for monitoring binding of serum antibodies to fusion  
CC proteins and as an immunogen to induce and/or enhance immune responses.  
CC The coding sequences can be ligated with a coding sequence of another  
CC molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and  
CC can be used in vivo as a DNA vaccine. This polynucleotide sequence  
CC represents the DNA encoding the Ra12-mammaglobin fusion protein.  
XX  
XX Sequence 672 BP; 160 A; 185 C; 187 G; 140 T; 0 other;

Alignment Scores:  
Pred. No.: 1.46e-14 Length: 672  
Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 23 Gaps: 0

US-09-684-215A-17 (1-30) x AAL40772 (1-672)

QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProile 20  
DB 25 ACGGCCGCGTCCGATACCTCCAGCTGCCAGGGTGGCGAGGATTCCGCATTCGCATC 84  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30  
DB 85 GGGCAGCGATGGCGATCGCGGCCAGATC 114

Search completed: September 5, 2003, 09:04:17  
Job time : 59.5747 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 5, 2003, 08:54:40 ; Search time 191.628 Seconds  
(without alignments)  
1584.308 Million cell updates/sec

Title: US-09-684-215A-4

Perfect score: 675

Sequence: 1 TAASDNFQLSQGGQGFAPFI.....QTKSGGTRGNTVLAEGPPA 132

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1537136 seqs, 1149988732 residues

Total number of hits satisfying chosen parameters: 3074272

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-O/cgn2\_1/USPTO.spool.p/US09684215/runat\_05092003\_072203\_619/app\_query.fasta\_1.853  
-DB=PublishedApplications\_NA\_QFAST-fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsmb2  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US09684215\_@CGN1.1.130\_@runat\_05092003\_072203\_619  
-NCPU=6 -ICPU=3 -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100  
-LOGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications\_NA:\*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
16: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
17: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	675	100.0	702	9	US-09-287-849-27 Sequence 27, Appl

2	675	100.0	702	12	US-10-359-460-27 Sequence 27, Appl
3	675	100.0	1068	10	US-09-712-363-15 Sequence 15, Appl
4	675	100.0	2287	9	US-09-287-849-1 Sequence 1, Appli
5	675	100.0	2287	12	US-10-359-460-1 Sequence 4, Appli
6	670	99.3	447	12	US-10-084-843-4 Sequence 4, Appli
7	670	99.3	447	12	US-10-193-002-4 Sequence 4, Appli
8	670	99.3	1872	12	US-10-084-843-17 Sequence 17, Appl
9	670	99.3	1872	12	US-10-193-002-17 Sequence 17, Appl
10	632	93.6	675	9	US-09-759-143-822 Sequence 822, App
11	632	93.6	675	9	US-09-780-669-822 Sequence 822, App
12	632	93.6	675	9	US-09-822-827-822 Sequence 822, App
13	632	93.6	675	10	US-09-895-793-822 Sequence 822, App
14	632	93.6	675	10	US-09-895-814-822 Sequence 822, App
15	632	93.6	675	12	US-10-144-678A-822 Sequence 822, App
16	632	93.6	822	13	US-10-012-896-822 Sequence 1862, Ap
17	632	93.6	822	10	US-09-736-457-1862 Sequence 1862, Ap
18	632	93.6	822	10	US-09-902-941-1862 Sequence 1862, Ap
19	632	93.6	822	10	US-09-849-626-1862 Sequence 1862, Ap
20	632	93.6	822	14	US-10-017-754-1862 Sequence 353, App
21	632	93.6	900	9	US-09-735-705-353 Sequence 353, App
22	632	93.6	900	10	US-09-850-716A-353 Sequence 353, App
23	632	93.6	900	10	US-09-897-778-353 Sequence 353, App
24	632	93.6	900	12	US-10-117-982-353 Sequence 353, App
25	632	93.6	915	9	US-09-759-143-834 Sequence 834, App
26	632	93.6	915	9	US-09-780-669-834 Sequence 834, App
27	632	93.6	915	9	US-09-822-827-834 Sequence 834, App
28	632	93.6	915	10	US-09-895-793-834 Sequence 834, App
29	632	93.6	915	10	US-09-895-814-834 Sequence 834, App
30	632	93.6	915	12	US-10-144-678A-834 Sequence 834, App
31	632	93.6	915	13	US-10-012-896-834 Sequence 1861, Ap
32	632	93.6	945	10	US-09-736-457-1861 Sequence 1861, Ap
33	632	93.6	945	10	US-09-902-941-1861 Sequence 1861, Ap
34	632	93.6	945	10	US-09-849-626-1861 Sequence 1861, Ap
35	632	93.6	945	14	US-10-017-754-1861 Sequence 351, App
36	632	93.6	1012	9	US-09-735-705-351 Sequence 351, App
37	632	93.6	1012	10	US-09-850-716A-351 Sequence 351, App
38	632	93.6	1012	10	US-09-897-778-351 Sequence 351, App
39	632	93.6	1012	12	US-10-117-982-351 Sequence 1084, Ap
40	632	93.6	1035	9	US-09-922-217-1084 Sequence 1084, Ap
41	632	93.6	1035	10	US-09-833-263-1084 Sequence 1084, Ap
42	632	93.6	1035	11	US-09-938-864-388 Sequence 388, App
43	632	93.6	1035	13	US-10-025-380-1084 Sequence 1084, Ap
44	632	93.6	1035	14	US-10-125-635A-388 Sequence 388, App
45	632	93.6	1035	14	US-10-002-603-388 Sequence 388, App

## ALIGNMENTS

### RESULT 1

US-09-287-849-27

; Sequence 27, Application US/09287849

; Patent No. US2002009459A1

; GENERAL INFORMATION:

; APPLICANT: Skeiky, Steven G.

; APPLICANT: Reedy, David C.

; APPLICANT: Dillion, David C.

; APPLICANT: Alderson, Mark

; APPLICANT: Campos-Neto, Antonio

; APPLICANT: Corixa Corporation

; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens

; TITLE OF INVENTION: and Their Uses

; FILE REFERENCE: 014058-00902005

; CURRENT APPLICATION NUMBER: US/09/287,849

; CURRENT FILING DATE: 1999-04-07

; PRIOR APPLICATION NUMBER: US 08/818,112

; PRIOR FILING DATE: 1997-03-13

; PRIOR APPLICATION NUMBER: US 08/942,578

; PRIOR FILING DATE: 1997-10-01

; PRIOR APPLICATION NUMBER: US 09/025,197

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 09/056,556

; PRIOR FILING DATE: 1998-04-07

; PRIOR APPLICATION NUMBER: US 09/223,040

```
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 702
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:bi-fusion
; OTHER INFORMATION: protein Ral2-DPPD (designated Mtb24), reading
; OTHER INFORMATION: frame 1
; NAME/KEY: CDS
; LOCATION: (1)..(693)
; OTHER INFORMATION: bi-fusion protein Ral2-DPPD (designated Mtb24),
; OTHER INFORMATION: reading frame 1
; NAME/KEY: CDS
; LOCATION: (2)..(700)
; OTHER INFORMATION: reading frame 2
; NAME/KEY: CDS
; LOCATION: (3)..(701)
; OTHER INFORMATION: reading frame 3
US-09-287-849-27

Alignment Scores:
Pred. No.: 2,266-69 Length: 702
Score: 675.00 Matches: 132
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-684-215A-4 (1-132) x US-09-287-849-27 (1-702)
QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20
DB 25 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGTGGCAGGAGTTCGCCATTCGGATC 84
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40
DB 85 GGGCAGGCGATGGCGATCGCGGCCAGATCCGATCGGTGGGGTCAACCCACCGTTTCA 144
QY 41 IleGlyProThrAlaPheLeuGlyValValAlaAspAsnGlyAsnGlyAlaArg 60
DB 145 ATCGGGCTACCGCTTCTCGGTGGGTGTGTCGACAAACACGCGACGGCCACGA 204
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
DB 205 GTCCAAACGCGTGGTGGGAGCGCTCCGCGGCGCAAGTCTCGCATCTCCACCGCGCGT 264
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
DB 265 ATCCACCGCGTGGTGGGAGCGCTCCGCGGCGCAAGTCTCGCATCTCCACCGCGCGT 324
QY 101 GlyHisHsProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
DB 325 GGGCATATATCCCGTGAGTCATCTCGGTGACCTCGCAACCAACGTCGGGGCGACGGGT 384
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
DB 385 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 420

RESULT 2
US-10-359-460-27
; Sequence 27, Application US/10359460
; Publication No. US20030147911A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
```

Db	910	ATCACCGCGGTGCACGGCGGTCCGATCAACTCGGCCACCGCGATGGCGGACGGCGCTTAAAC	969
Qy	101	GlyHisHisProGlyAspValIleSerValThrTrrpGlnThrLysSerGlyGlyThrArg	120
Db	970	GGCATCATCCCGGTGACGTCACTCGGTGACCTGGCAACCAAGTCGGGGGCGCACCGCT	1029
Qy	121	ThrGlyAsnValThrLeuAlaGluGlyProProAla	132
Db	1030	ACAGGGAACGTGACATTGGCCGAGGGACCCCGGCC	1065
RESULT 4			
US-09-287-849-1			
; Sequence 1, Application US/09287849			
; Patent No. US20020009459A1			
; GENERAL INFORMATION:			
; APPLICANT: Reed, Steven G.			
; APPLICANT: Skeiky, Yasir A.W.			
; APPLICANT: Dillon, Davin C.			
; APPLICANT: Alderson, Mark			
; APPLICANT: Campos-Neto, Antonio			
; APPLICANT: Corixa Corporation			
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens			
; TITLE OF INVENTION: and Their Uses			
; FILE REFERENCE: 014058-00902005			
; CURRENT APPLICATION NUMBER: US/09/287,849			
; CURRENT FILING DATE: 1999-04-07			
; PRIOR APPLICATION NUMBER: US 08/818,112			
; PRIOR FILING DATE: 1997-03-13			
; PRIOR APPLICATION NUMBER: US 08/942,578			
; PRIOR FILING DATE: 1997-10-01			
; PRIOR APPLICATION NUMBER: US 09/025,197			
; PRIOR FILING DATE: 1998-02-18			
; PRIOR APPLICATION NUMBER: US 09/056,556			
; PRIOR FILING DATE: 1998-04-07			
; PRIOR APPLICATION NUMBER: US 09/223,040			
; PRIOR FILING DATE: 1998-12-30			
; NUMBER OF SEQ ID NOS: 46			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 1			
; LENGTH: 2287			
; TYPE: DNA			
; ORGANISM: Artificial Sequence			
; FEATURE:			
; OTHER INFORMATION: Description of Artificial Sequence:tri-fusion			
; OTHER INFORMATION: protein Ra12-TbH9-Ra35 (designated Mtb32-Mtb39			
; OTHER INFORMATION: fusion)			
; NAME/KEY: modified_base			
; LOCATION: (30)			
; OTHER INFORMATION: n = g, a, c or t			
; NAME/KEY: modified_base			
; LOCATION: (33)			
; OTHER INFORMATION: n = g, a, c or t			
; NAME/KEY: CDS			
; LOCATION: (42)..(2231)			
; NAME/KEY: modified_base			
; LOCATION: (2270)			
; OTHER INFORMATION: n = g, a, c or t			
US-09-287-849-1			
Alignment Scores:			
Pred. No.:			
Score:			
Percent Similarity:			
Best Local Similarity:			
Query Match:			
DB:			
US-09-684-215A-4 (1-132) x US-09-287-849-1 (1-2287)			
Qy	1	ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle	20
Db	63	ACGGCGCGGTCCGTAACATTCCAGCTGCCAGGGTGGCAGGGAATTCGCCATTCCGATC	122

```
OY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
Db 123 GGGCAGCGCATCGCGGCGCAGATCCGATCGGTGGGGGTCAACCCACCGTTCAT 182
OY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArg 60
Db 183 ATCGGGCCCTACCGCCCTTCCTCGGCTGGGTGTGTGTCGACAAACGCGCAACGGCGCAGCA 242
OY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
Db 243 GTCCACAGCGGTGGTGGGAGCGCTCGGGCGGCAAGTCTCGGCATCTCCACCGCGGACGTG 302
OY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
Db 303 ATCACCGCGGTGACGCGCTCCGATCAACTCGGCGCACCGCATCGGCGGCGCGCTTAAC 362
OY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrIlysserGlyThrArg 120
Db 363 GGGCATCATCCCGGTGACGTCTATCTCGGTGACCTGGCAAAACCAAGTCGGGCGGCGCGGT 422
OY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
Db 423 ACAGGAGCGTGACATTGGCCGAGGAGGCCCGCGGCC 458
```

## RESULT 5

```
US-10-359-460-1
; Sequence 1, Application US/10359460
; Publication No. US20030147911A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
; TITLE OF INVENTION: and their Uses
; FILE REFERENCE: 014058-009020US
; CURRENT APPLICATION NUMBER: US/10/359,460
; CURRENT FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: US/09/287,849
; PRIOR FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: US 08/818,112
; PRIOR FILING DATE: 1997-03-13
; PRIOR APPLICATION NUMBER: US 08/942,578
; PRIOR FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: US 09/025,197
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 09/056,556
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2287
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: tri-fusion
; OTHER INFORMATION: protein Ra12-TbH9-Ra35 (designated Mtb32-Mtb39
; OTHER INFORMATION: fusion)
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (30)
; OTHER INFORMATION: n = g, a, c or t
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (33)
; OTHER INFORMATION: n = g, a, c or t
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (42)..(2231)
```

```
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (2270)
; OTHER INFORMATION: n = g, a, c or t
US-10-359-460-1
Alignment Scores:
Pred. No.: 9,31e-69 Length: 2287
Score: 675.00 Matches: 132
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0
US-09-684-215A-4 (1-132) x US-10-359-460-1 (1-2287)
OY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProfile 20
Db 63 ACGGCGCGGTTCGATAAATTCAGCTGTCCAGGGTGGCGAGGGATTCGCCATTCGGATC 122
OY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
Db 123 GGGCAGCGCATCGCGATCGCGGCGCAGATCCGATCGGTGGGGGTCAACCCACCGTTCAT 182
OY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArg 60
Db 183 ATCGGCGCTACCGCCCTTCCTCGGCTGGGTGTGTGTCGACAAACGCGCAACGGCGCAGCA 242
OY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
Db 243 GTCCACAGCGGTGGTGGGAGCGCTCGGGCGGCAAGTCTCGGCATCTCCACCGCGGACGTG 302
OY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
Db 303 ATCACCGCGGTGCGAGCGCTCCGATCAACTCGGCGCACCGCATGGCGGCGCGCTTAAC 362
OY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrIlysserGlyThrArg 120
Db 363 GGGCATCATCGCGGTGACGTCTATCTCGGTGACCTGGCAAAACCAAGTCGGGCGGCGCGGT 422
OY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
Db 423 ACAGGAGCGTGACATTGGCCGAGGAGGCCCGCGGCC 458
RESULT 6
US-10-084-843-4
; Sequence 4, Application US/10084843
; Publication No. US20030143243A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Houghton, Raymond
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Twardzik, Daniel R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Hendrickson, Ronald C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
; AND DIAGNOSIS OF TUBERCULOSIS
; NUMBER OF SEQUENCES: 355
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```



;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/10/084,843  
;; FILING DATE: 25-Feb-2002  
;; CLASSIFICATION: <Unknown>  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/072,967  
;; FILING DATE: 05-MAY-1998  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Makl, David J.  
;; REGISTRATION NUMBER: 31,392  
;; REFERENCE/DOCKET NUMBER: 210121.411C9  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (206) 622-4900  
;; TELEFAX: (206) 682-6031  
;; INFORMATION FOR SEQ ID NO: 4:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 447 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-10-084-843-4

Alignment Scores:  
Pred. No.: 5,01e-69 Length: 447  
Score: 670.00 Matches: 131  
Percent Similarity: 99.24% Conservative: 0  
Best Local Similarity: 99.24% Mismatches: 1  
Query Match: 99.26% Indels: 0  
DB: 12 Gaps: 0

US-09-684-215A-4 (1-132) x US-10-084-843-4 (1-447)

QY 1 ThrAlaAlaSerAspAsnPhelGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
Db 11 ACGGCGCGTCCGATAAATCCAGCTGTCCAGGTGGCAGGATTCGCATTCGCATC 70  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40  
Db 71 GGGCAGCGATGGCGATCGCGGCCAGATCCGATCGGTGGGGGTCAACCCACCGTTCAT 130  
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAlaArg 60  
Db 131 ATCGGCGCTACCGCTTCCTCGGTGGGTGTGTCACAAACGCGACGCGCGACCA 190  
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 191 GTCCACGCGTGGTGGGAGCGCTCCGATCAACTCCGCGCAAGTCTCGGCATCTCCACCGCGACGTG 250  
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 251 ATCACCGCGTCCAGCGCTCCGATCAACTCCGCGCAAGTCTCGGCATCTCCACCGCGACGTG 310  
QY 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
Db 311 GGCATCATCCCGTACGCTCATCTCGGTGAACGGCAACCAAGTCGGGCGGCGACGCGT 370  
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
Db 371 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 406

## RESULT 7

US-10-193-002-4  
; Sequence 4, Application US/10193002  
; Publication No. US20030135026A1

## GENERAL INFORMATION:

;; APPLICANT: Reed, Steven G.  
;; Skelky, Yasir A.W.  
;; Dillon, Davin C.  
;; Campos-Neto, Antonia  
;; Houghton, Raymond  
;; Vedrick, Thomas S.  
;; Twardzik, Daniel R.

;; Lodes, Michael J.  
;; Hendrickson, Ronald C.  
;; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF  
;; TUBERCULOSIS  
;; NUMBER OF SEQUENCES: 350  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: SEED and BERRY LLP  
;; STREET: 6300 Columbia Center, 701 Fifth Avenue  
;; CITY: Seattle  
;; STATE: Washington  
;; COUNTRY: USA  
;; ZIP: 98104-7092  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA: US/10/193,002  
;; FILING DATE: 10-Jul-2002  
;; CLASSIFICATION: <Unknown>  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/072,596  
;; FILING DATE: 05-MAY-1998  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Makl, David J.  
;; REGISTRATION NUMBER: 31,392  
;; REFERENCE/DOCKET NUMBER: 210121.417C9  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (206) 622-4900  
;; TELEFAX: (206) 682-6031  
;; INFORMATION FOR SEQ ID NO: 4:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 447 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-10-193-002-4

Alignment Scores:  
Pred. No.: 5,01e-69 Length: 447  
Score: 670.00 Matches: 131  
Percent Similarity: 99.24% Conservative: 0  
Best Local Similarity: 99.24% Mismatches: 1  
Query Match: 99.26% Indels: 0  
DB: 12 Gaps: 0

US-09-684-215A-4 (1-132) x US-10-193-002-4 (1-447)  
QY 1 ThrAlaAlaSerAspAsnPhelGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
Db 11 ACGGCGCGTCCGATAAATCCAGCTGTCCAGGTGGCAGGATTCGCATTCGCATC 70  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40  
Db 71 GGGCAGCGATGGCGATCGCGGCCAGATCCGATCGGTGGGGGTCAACCCACCGTTCAT 130  
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAlaArg 60  
Db 131 ATCGGCGCTACCGCTTCCTCGGTGGGTGTGTCACAAACGCGACGCGCGACCA 190  
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 191 GTCCACGCGTGGTGGGAGCGCTCCGATCAACTCCGCGCAAGTCTCGGCATCTCCACCGCGACGTG 250  
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 251 ATCACCGCGTCCAGCGCTCCGATCAACTCCGCGCAAGTCTCGGCATCTCCACCGCGACGTG 310  
QY 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
Db 311 GGCATCATCCCGTACGCTCATCTCGGTGAACGGCAACCAAGTCGGGCGGCGACGCGT 370

QY	121	ThrGlyAsnValThrLeuAlaGluGlyProProAla 132	
Db	371	ACAGGGAACGTGACATTCGGCGAGGAGCCCGCGCC 406	
RESULT 8			
US-10-084-843-17			
; Sequence 17, Application US/10084843			
; Publication No. US20030143243A1			
; GENERAL INFORMATION:			
; APPLICANT: Reed, Steven G.			
; Skelky, Yasir A.W.			
; Dillon, Davin C.			
; Campos-Neto, Antonio			
; Houghton, Raymond			
; Vedvick, Thomas S.			
; Twardzik, Daniel R.			
; Lodes, Michael J.			
; Hendrickson, Ronald C.			
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY			
; AND DIAGNOSIS OF TUBERCULOSIS			
; NUMBER OF SEQUENCES: 355			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: SEED AND BERRY LLP			
; STREET: 6300 Columbia Center, 701 Fifth Avenue			
; CITY: Seattle			
; STATE: Washington			
; COUNTRY: USA			
; ZIP: 98104-7092			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: PatentIn Release #1.0, Version #1.30			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/10/084,843			
; FILING DATE: 25-Feb-2002			
; CLASSIFICATION: <Unknown>			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: US/09/072,967			
; FILING DATE: 05-MAY-1998			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Maki, David J.			
; REGISTRATION NUMBER: 31,392			
; REFERENCE/DOCKET NUMBER: 210121.411C9			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: (206) 622-4900			
; TELEFAX: (206) 682-6031			
; INFORMATION FOR SEQ ID NO: 17:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 1872 base pairs			
; TYPE: nucleic acid			
; STRANDEDNESS: single			
; TOPOLOGY: linear			
; SEQUENCE DESCRIPTION: SEQ ID NO: 17:			
US-10-084-843-17			
Alignment Scores:			
Pred. No.: 2,79e-68 Length: 1872			
Score: 670.00 Matches: 131			
Percent Similarity: 99.24% Conservative: 0			
Best Local Similarity: 99.24% Mismatches: 1			
Query Match: 99.26% Indels: 0			
DB: 12 Gaps: 0			
US-09-684-215A-4 (1-132) x US-10-084-843-17 (1-1872)			
QY	1	ThrAlaAlaSerAspAspGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20	
Db	758	ACGGCGCGTCGGATACCTCCAGCTGTCACAGGTGGCGAGGATTCGGCATTCGGATC 817	
QY	21	GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40	

```
Alignment Scores:
Pred. No.: 2,79e-68 Length: 1872
Score: 670.00 Matches: 131
Percent Similarity: 99.24% Conservative: 0
Best Local Similarity: 99.24% Mismatches: 1
Query Match: 99.26% Indels: 0
DB: 12 Gaps: 0

US-09-684-215A-4 (1-132) x US-10-193-002-17 (1-1872)

Qy 1 ThrAlaAlaSerAspAsnPhelGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20
Db 758 ACGCCGCGTCCGATAACTTCAGCTGTCCAGGGTGGGAGGATTGCCATTCCGATC 817
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
Db 818 GGGCAGGCGATGGCGATCGCGGCAATCCGATCGGGTGGGGGTCAACCCCGTTCAT 877
Qy 41 IleGlyProThrAlaPheLeuGlyValValAlaAspAsnAsnGlyValAlaArg 60
Db 878 ATCGGGCTACCGCTTCCTCGGCTTGGTGTTCGACAAACAACGGCAACGGCGCACGA 937
Qy 61 ValGlnArgValValGlySerAlaProAlaIleSerLeuGlyIleSerThrGlyAspVal 80
Db 938 GTCCAACGCGTGGTGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGGCGACGTG 997
Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAlaAspAlaLeuAsn 100
Db 998 ATCAGCGGGTGCAGGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGGCTTAAC 1057
Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
Db 1058 GGGCATCATCCGCTGACGTCTCATCTCGGTGAACCTGGCAACCAAGTCGGGGCGGCGCGT 1117
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
Db 1118 ACAGGGAACGTGACATTTGGCGGAGGAGCCCGCGGCC 1153

RESULT 10
; Sequence 822, Application US/09759143
; Patent No. US2002002248A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, David C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C23
; CURRENT APPLICATION NUMBER: US/09/759,143
; CURRENT FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 822
; LENGTH: 675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-759-143-822

Alignment Scores:
Pred. No.: 2.1e-64 Length: 675
Score: 632.00 Matches: 126
Percent Similarity: 96.21% Conservative: 1
Best Local Similarity: 95.45% Mismatches: 1
Query Match: 93.63% Indels: 4
DB: 9 Gaps: 1

US-09-684-215A-4 (1-132) x US-09-759-143-822 (1-675)

Qy 1 ThrAlaAlaSerAspAsnPhelGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20
Db 22 ACGCCGCGTCCGATAACTTCAGCTGTCCAGGGTGGGAGGATTGCCATTCCGATC 81
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
Db 82 GGGCAGGCGATGGCGATCGCGGGCAGATCAAG-----CTTCCCGCCGTTCT 129
Qy 41 IleGlyProThrAlaPheLeuGlyValValAlaAspAsnAsnGlyValAlaArg 60
Db 130 ATCGGGCTACCGCTTCCTCGGCTTGGTGTTCGACAAACAACGGCAACGGCGCACGA 189
Qy 61 ValGlnArgValValGlySerAlaProAlaIleSerLeuGlyIleSerThrGlyAspVal 80
Db 190 GTCCAACGCGTGGTGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGGCGACGTG 249
Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAlaAspAlaLeuAsn 100
Db* 250 ATCAGCGGGTGCAGGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGGCTTAAC 309
Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
Db 310 GGGCATCATCCGCTGACGTCTCATCTCGGTGAACCTGGCAACCAAGTCGGGGCGGCGCGT 369
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
Db 370 ACAGGGAACGTGACATTTGGCGGAGGAGCCCGCGGCC 405

RESULT 11
US-09-780-669-822
; Sequence 822, Application US/09780669
; Patent No. US2002005197A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, David C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C24
; CURRENT APPLICATION NUMBER: US/09/780,669
; CURRENT FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 943
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 822
; LENGTH: 675
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-780-669-822

Alignment Scores:
Pred. No.: 2,1e-64 Length: 675
Score: 632.00 Matches: 126
Percent Similarity: 96.21% Conservative: 1
Best Local Similarity: 95.45% Mismatches: 1
Query Match: 93.63% Indels: 4
DB: 9 Gaps: 1

US-09-684-215A-4 (1-132) x US-09-780-669-822 (1-675)

QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20
|||||
Db 22 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGGTGGCAGGATTCGCCATTCCGATC 81
|||||
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHls 40
|||||
Db 82 GGCAGGCGATCGCGATTCAGCTGTCCAGGTGGCAGGATTCGCCATTCCGATC 129
|||||
QY 41 IleGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHls 60
|||||
Db 130 ATCGGCGCTACCGCTTCCGCTTGGGTGTGTCGACAAACGCGCAGCAGCA 189
|||||
QY 61 ValGlnArgValValcIlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
|||||
Db 190 GTCCAAACGCGTGGTGGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGCGCAGCG 249
|||||
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAlaLeuAsn 100
|||||
Db 250 ATCACCAGCGGTGCGAGCGGCTCCGATCAACTCGGCCACCGCGATGGCGGCGCTTAAC 309
|||||
QY 101 GlyHlsHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
|||||
Db 310 GGCATCATCCCGGTGAGCTCATCTCGGTGACCTGGCAACCAAGTCGGCGGCGCAGCGT 369
|||||
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
|||||
Db 370 ACAGGAAACGTGACATTGGCGGAGGACCCCGGCC 405
|||||

RESULT 13
US-09-895-793-822
; Sequence 822, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darriek
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.534C2
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 822
; LENGTH: 675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-895-793-822

Alignment Scores:
Pred. No.: 2,1e-64 Length: 675
Score: 632.00 Matches: 126
Percent Similarity: 96.21% Conservative: 1
Best Local Similarity: 95.45% Mismatches: 1
Query Match: 93.63% Indels: 4
DB: 9 Gaps: 1

US-09-684-215A-4 (1-132) x US-09-822-827-822 (1-675)

QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20
|||||
Db 22 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGTGGCAGGATTCGCCATTCCGATC 81
|||||
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHls 40
|||||
Db 82 GGCAGGCGATCGCGATTCAGCTGTCCAGGTGGCAGGATTCGCCATTCCGATC 129
|||||
QY 41 IleGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHls 60
|||||
Db 130 ATCGGCGCTACCGCTTCCGCTTGGGTGTGTCGACAAACGCGCAGCAGCA 189
|||||
QY 61 ValGlnArgValValcIlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
|||||
Db 190 GTCCAAACGCGTGGTGGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGCGCAGCG 249
|||||
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAlaLeuAsn 100
|||||
Db 250 ATCACCAGCGGTGCGAGCGGCTCCGATCAACTCGGCCACCGCGATGGCGGCGCTTAAC 309
|||||
QY 101 GlyHlsHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
|||||
Db 310 GGCATCATCCCGGTGAGCTCATCTCGGTGACCTGGCAACCAAGTCGGCGGCGCAGCGT 369
|||||
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
|||||
Db 370 ACAGGAAACGTGACATTGGCGGAGGACCCCGGCC 405
|||||

RESULT 12
US-09-822-827-822
; Sequence 822, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.534C1
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 822
; LENGTH: 675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-822-827-822

Alignment Scores:
Pred. No.: 2,1e-64 Length: 675
Score: 632.00 Matches: 126
Percent Similarity: 96.21% Conservative: 1
Best Local Similarity: 95.45% Mismatches: 1
Query Match: 93.63% Indels: 4
DB: 9 Gaps: 1

US-09-684-215A-4 (1-132) x US-09-822-827-822 (1-675)

QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20
|||||
```

```
US-09-684-215A-4 (1-132) x US-09-895-793-822 (1-675)
QY 1 ThrAlaAlaSerAspAsnPhGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 22 ACCGCCGCGTCCGATACCTCCAGCTGTCCAGGGTGGGAGGATTCGGCATTCGCATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
Db 82 GGGCAGCGGATGCGATCGCGGCCAGATCAAG-----CTTCCACCGTTTCAT 129
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
Db 130 ATCGGGCCTACCCCTTCCTCGGCTTGGGTGTTCGACAAACACGCGCACGACGA 189
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
Db 190 GTCCACGCGTGTGCGGAGCGTCCGGCGGCAAGTCTCGCATCTCCACCGCGAGTG 249
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
Db 250 ATCAGCGCGTGCAGCGGCTCGGATCAACTCGGCCACCGCATGGCGGCGCACGCGT 309
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
Db 310 GGGCATCATCCCGGTGACGTCTCGGTGACCTGGCAACCAAGTGGCGGCGCACGCGT 369
QY 121 ThrGlyAsnValThrLeuAlaGlyProProAla 132
Db 370 ACAGGGAACGTGACATGGCGGAGGACCCCGGCC 405

RESULT 14
US-09-895-814-822
; Sequence 822, Application US/09895814
; Publication No. US20020193296A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C26
; CURRENT APPLICATION NUMBER: US/09/895,814
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 990
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 822
; LENGTH: 675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-895-814-822
Alignment Scores: 2.1e-64 Length: 675
Pred. No.: 2.1e-64 Length: 675
```

```
Score: 632.00 Matches: 126
Percent Similarity: 96.21% Conservative: 1
Best Local Similarity: 95.45% Mismatches: 1
Query Match: 93.63% Indels: 4
Db: 10 Gaps: 1

US-09-684-215A-4 (1-132) x US-09-895-814-822 (1-675)
QY 1 ThrAlaAlaSerAspAsnPhGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 22 ACCGCCGCGTCCGATACCTCCAGCTGTCCAGGGTGGGAGGATTCGGCATTCGCATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
Db 82 GGGCAGCGGATGCGATCGCGGCCAGATCAAG-----CTTCCACCGTTTCAT 129
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
Db 130 ATCGGGCCTACCCCTTCCTCGGCTTGGGTGTTCGACAAACACGCGCACGACGA 189
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
Db 190 GTCCACGCGTGTGCGGAGCGTCCGGCGGCAAGTCTCGCATCTCCACCGCGAGTG 249
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
Db 250 ATCAGCGCGTGCAGCGGCTCGGATCAACTCGGCCACCGCATGGCGGCGCACGCGT 309
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
Db 310 GGGCATCATCCCGGTGACGTCTCGGTGACCTGGCAACCAAGTGGCGGCGCACGCGT 369
QY 121 ThrGlyAsnValThrLeuAlaGlyProProAla 132
Db 370 ACAGGGAACGTGACATGGCGGAGGACCCCGGCC 405

RESULT 15
US-10-144-678A-822
; Sequence 822, Application US/10144678A
; Publication No. US20030157089A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Hepler, William T.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals y de Bassols, Carlota
; APPLICANT: Foy, Teresa M.
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Deng, Ta
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C28
; CURRENT APPLICATION NUMBER: US/10/144,678A
; CURRENT FILING DATE: 2002-08-12
; NUMBER OF SEQ ID NOS: 1033
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 822
```

; LENGTH: 675  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-144-678A-822

Alignment Scores:  
 Pred. No.: 2,1e-64 Length: 675  
 Score: 632.00 Matches: 126  
 Percent Similarity: 96.21% Conservative: 1  
 Best Local Similarity: 95.45% Mismatches: 1  
 Query Match: 93.63% Indels: 4  
 DB: 12 Gaps: 1

US-09-684-215A-4 (1-132) x US-10-144-678A-822 (1-675)

QY	1	ThrAlaAlaSerAspGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle	20
DB	22	ACGGCGCGTCCGATCACTCCAGCTGTCCAGGTGGCAGGGATTGCCATTCCGATC	81
QY	21	GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis	40
DB	82	GGGCAGGGCATGGCGATCGGGGCCAGATCAAG-----CTTCCACCGTTCAT	129
QY	41	IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAlaArg	60
DB	130	ATCGGGCCTACCGCTTCTCGGCTTGGGTGTGTGACAAACGCAACGGCGCAGCA	189
QY	61	ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal	80
DB	190	GTCCAACGGTGGTGGGAGCGCTCCGGGGGCAAGTCTCGGCATCTCCACGGCGACGTG	249
QY	81	IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn	100
DB	250	ATCACCGCGTCCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGACGCGCTTAAC	309
QY	101	GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg	120
DB	310	GGGCATCATCCCGGTGACGTCTCTCGGTGACCTGGCAACCAAGTCGGGGCGCACGGT	369
QY	121	ThrGlyAsnValThrLeuAlaGluGlyProProAla	132
DB	370	ACAGGGAACGTGACATTGGCCGAGGACCCCGGCC	405

Search completed: September 5, 2003, 12:04:43  
 Job time : 196.961 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 5, 2003, 07:24:34 ; Search time 248.497 Seconds  
(without alignments)  
1390.474 Million cell updates/sec

Title: US-09-684-215A-18

Perfect score: 653

Sequence: 1 TRAASNFLSQSGGQFAIPI.....QTKSGGTRGNTVLRGPPA 128

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Xgapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-O=Csq2\_1/USPTO.spool\_p/US09684215/runat\_05092003\_072201\_548/app\_query.fasta\_1.853  
-DB=N\_Geneseq\_19Jun03 -QFAST=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USPR=US09684215 -CGN\_1.1.626 @runat\_05092003\_072201\_548 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*
- 2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*
- 3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*
- 4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*
- 5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*
- 6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*
- 7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*
- 8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*
- 9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*
- 10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*
- 11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*
- 12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*
- 13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*
- 14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*
- 15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*
- 16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*
- 17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*
- 18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*
- 19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*
- 20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*
- 21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*
- 22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*
- 23: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*
- 24: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*
- 25: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	653	100.0	675	22	AAS64132	Human /M. tubercul
2	653	100.0	675	22	AH93896	Ral2-P510S-C const
3	653	100.0	675	24	ABL95503	Ral2- P510S-C cons
4	653	100.0	675	25	ACA59940	Prostate cancer th
5	653	100.0	822	24	ABK39769	DNA encoding lung
6	653	100.0	822	25	ACA12098	Human lung cancer
7	653	100.0	822	25	ACA03284	Lung cancer therap
8	653	100.0	894	24	ABK27798	Human colon cancer
9	653	100.0	900	24	ABO92443	Human lung cancer
10	653	100.0	900	24	ABL49257	Ral2/C-terminal po
11	653	100.0	915	22	AAS64141	Human /M. tubercul
12	653	100.0	915	22	AH93905	Ral2-P775P-ORF3 c
13	653	100.0	915	24	ABL95512	Ral2- P775P-ORF3 c
14	653	100.0	915	25	ACA59949	Prostate cancer th
15	653	100.0	945	24	ABK39768	DNA encoding lung
16	653	100.0	945	25	ACA12097	Human lung cancer
17	653	100.0	945	25	ACA03283	Lung cancer therap
18	653	100.0	1012	24	ABO92442	Human lung cancer
19	653	100.0	1012	24	ABL49256	Ral2/N-terminal po
20	653	100.0	1035	24	ABK69714	DNA encoding human
21	653	100.0	1035	25	ABK33699	Human colon tumour
22	653	100.0	1155	24	ABK39775	DNA encoding Ral2-
23	653	100.0	1155	25	ACA12104	Human Ral2-l985P f
24	653	100.0	1155	25	ACA03290	Lung cancer therap
25	653	100.0	1203	22	AAS64153	Human /M. tubercul
26	653	100.0	1203	22	AH93917	Ral2-P501S-E2 cons
27	653	100.0	1203	24	ABL95524	Ral2-P501S-E2 cons
28	653	100.0	1203	25	ACA59961	Prostate cancer th
29	653	100.0	1263	24	ABK69715	DNA encoding human
30	653	100.0	1464	22	AH56353	DNA encoding Chlam
31	653	100.0	1464	24	ABL92582	Chlamydia trachoma
32	653	100.0	1557	22	AH56341	DNA encoding Chlam
33	653	100.0	1557	24	ABL92570	Chlamydia trachoma
34	653	100.0	1578	21	AAAG4764	C. pneumoniae sero
35	653	100.0	1578	22	AH56267	Chlamydia trachoma
36	653	100.0	1578	24	ABL92496	Chlamydia DNA sequ
37	653	100.0	1590	24	AH599869	Breast tumour-spec
38	653	100.0	1707	24	ABK69716	DNA encoding human
39	653	100.0	1752	22	AH56356	DNA encoding Chlam
40	653	100.0	1752	24	ABL92585	Chlamydia trachoma
41	653	100.0	1758	22	AH56344	DNA encoding Chlam
42	653	100.0	1758	24	ABL92573	Chlamydia trachoma
43	653	100.0	1758	25	ACA59997	Prostate cancer th
44	653	100.0	1860	22	AH56323	DNA encoding Chlam
45	653	100.0	1860	24	ABL92552	Chlamydia trachoma

ALIGNMENTS

RESULT 1  
AAS64132

ID AAS64132 standard; cDNA; 675 BP.

XX AAS64132;

AC AAS64132;

DT 29-JAN-2002 (first entry)

XX Human /M. tuberculosis Ral2 fusion protein Ral2-P510S-C cDNA.

XX Human; prostate cancer; ss; cytostatic; immunostimulant; tumour.

XX Chimeric - Homo sapiens.

OS Chimeric - Microbacterium tuberculosis.

OS Synthetic.

PN WO200173032-A2.

```

XX PD 04-OCT-2001.
XX PF 27-MAR-2001; 2001WO-US09919.
XX PR 27-MAR-2000; 2000US-0536857.
XX PR 09-MAY-2000; 2000US-0568100.
XX PR 12-MAY-2000; 2000US-0570737.
XX PR 13-JUN-2000; 2000US-0593793.
XX PR 27-JUN-2000; 2000US-0605783.
XX PR 10-AUG-2000; 2000US-0636215.
XX PR 29-AUG-2000; 2000US-0651236.
XX PR 06-SEP-2000; 2000US-0657279.
XX PR 02-OCT-2000; 2000US-0679426.
XX PR 10-OCT-2000; 2000US-0685166.
XX PA (CORI-) CORIXA CORP.
XX PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;
PI Fanger GR, Retter MW, Stoik JA, Day CH, Vedvick TS, Carter D;
PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;
XX WPI: 2001-639232/73.
XX DR P-PSDB; AA069899.
XX
XX New human prostate-specific polypeptides and polynucleotides useful for
XX the diagnosis and treatment of cancer, especially prostate cancer -
XX Example 17; Page 532; 579pp; English.
XX
XX The invention relates to isolated prostate-specific
XX polynucleotides, polypeptides, fusion proteins of the polypeptides,
XX antibodies raised against the polypeptides (or antigenic epitopes
XX derived from them) and antigen-presenting cells expressing the
XX polypeptides. The antibodies are useful for detecting the presence of
XX cancer, especially prostate cancer. The polypeptides, polynucleotides and
XX the antigen-presenting cells are useful for stimulating and/or expanding
XX T cells specific for a tumour protein, and for inhibiting the development
XX of cancer especially prostate cancer. Compositions comprising the
XX polynucleotide and/or polypeptide are useful for stimulating an immune
XX response, and for treating cancer. The oligonucleotide is useful for
XX detecting cancer. The present sequence is a prostate specific
XX polynucleotide of the invention.
XX
XX Sequence 675 BP; 162 A; 197 C; 190 G; 126 T; 0 other;

Alignment Scores:
Pred. No.: 5,71e-58 Length: 675
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 22 Gaps: 0

US-09-684-215a-18 (1-128) x AAS64132 (1-675)
QY 1 ThrAlaLeuSerAspAsnPhcGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
DB 22 ACGGCGCGTCCGATAACTCCAGCTGTCCAGGTGGCGAGGATTCGCCATTCGATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLeuProThrValHisIleGlyProThr 40
DB 82 GGGCAGGGCGATGGCGATCGGGGCGAGATCAAGCTCCAGCTTCATATCGGGCTACC 141
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
DB 142 GCCTTCCTCGGCTTGGGTGTGTCACACACACGCGACGCGCGACGATCCACCGCGTG 201
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
DB 202 GTCGGAGCGCTCCGCGCGCAAGTCTCGCATCTCCACCGCGCGAGCTGATCACCGCGTC 261
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100

```

---

```

DB 262 GACGGCGCTCGCATCACTCGCCACCGCGATGGCGGACGGCTTAACGGCGCATCATCCC 321
QY 101 GlyAspValIleSerValThrTrpClnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
DB 322 GGTGACGTCATCTCGGTGACCTGGCAACCAAGTCGGCGGACCGGTACAGGGGACGTCG 381
QY 121 ThrLeuAlaGlnGlyProProAla 128
DB 382 ACATTGGCGGAGGGGACCCCGGCC 405

RESULT 2
AAH93896
ID AAH93896 standard; cDNA; 675 BP.
XX
XX AAH93896;
XX
XX 04-OCT-2001 (first entry)
XX DE Ra12-PS10S-C construct cDNA sequence.
XX KW Human; prostate cancer; prostate-specific; diagnosis; vaccine;
XX KW cytostatic; gene therapy; metastasis; ss.
XX OS Homo sapiens.
XX PN WO200151633-A2.
XX PD 19-JUL-2001.
XX PF 16-JAN-2001; 2001WO-US01574.
XX PR 14-JAN-2000; 2000US-0483672.
XX PA (CORI-) CORIXA CORP.
XX PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SC;
PI Kalos MD, Fanger GR, Day CH, Retter MW, Stoik JA, Skeiky YAW;
PI Wang A, Meagher MJ;
XX DR WPI: 2001-425873/45.
XX PT New polynucleotide encoding a prostate-specific protein, for
XX PT diagnosing, monitoring and treating prostate cancer in a patient and
XX PT for use in vaccines.
XX PS Claim 8; Page 492-493; 543pp; English.
XX
XX The present invention describes polynucleotide sequences (I) which encode
XX prostate-specific proteins (II). (I) and (II) have cytostatic activity,
XX and can be used in vaccine production and gene therapy. (I), (II),
XX antibodies to (II), fusion proteins comprising (II), and isolated
XX T cells prepared using (I) or (II) are used to treat cancer in a patient.
XX (I) and the antibodies are also used in the detection of cancer in a
XX patient. The cancer that is diagnosed or treated is particularly
XX prostate cancer. (I) and (II) can be used in vaccines. The antibodies or
XX (I) can be used for monitoring the progression of cancer in a patient.
XX (I) and (II) can also be used to improve diagnostic and therapeutic
XX methods for prostate cancer. They can indicate the level of metastasis
XX as well as the prostate volume. AAH93357 to AAH93944 and AAH01115 to
XX AAH01318 represent polynucleotide and amino acid sequences used in the
XX exemplification of the present invention.
XX
XX Sequence 675 BP; 162 A; 197 C; 190 G; 126 T; 0 other;

Alignment Scores:
Pred. No.: 5,71e-58 Length: 675
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 22 Gaps: 0

```



US-09-684-215A-18 (1-128) x AAH93896 (1-675)

```
Qy 1 ThrAlaAlaSerAspAsnPhelGlnLeuSerGlnGlyGlnGlyPheAlaIleProile 20
Db 22 ACGCCGCGTCCGATAACTTCCAGCTGTCCAGGGTGGCAGGATTCCGATTCGGATC 81
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
Db 82 GGCAGGCGATGGCGATCGCGGCCAGATCAAGCTTCCACCGTTCATATCGGCGCTACC 141
Qy 41 AlaPheLeuGlyLeuGlyValValAlaAspAsnGlyAsnGlyAlaArgValGlnArgVal 60
Db 142 GCCTTCTCGGCTGGGTGGTGTTCGACAAACACGCGCAGCGGATCCACGCGGTG 201
Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
Db 202 GTCGGAGCGCTCGCGCGCAAGTCTCGCATCTCCACCGGCGAGTGCACGCGGTC 261
Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100
Db 262 GACGGCGCTCGATCAACTCGGCCACCGCATCGGCGGCGGCGCTTAACGGGCATCATCCC 321
Qy 101 GlyAspValIleSerValThrTpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
Db 322 GGTGACGTCACTCTCGGTGACCTGGCAACCAAGTCCGGCGGCGACGCGTACAGGGAACGTG 381
Qy 121 ThrLeuAlaGluGlyProProAla 128
Db 382 ACATTGGCGGAGGACCCCGGCC. 405
RESULT 3
ABL95503
ID ABL95503 standard; cDNA; 675 BP.
AC ABL95503;
XX
DT 19-JUL-2002 (first entry)
XX
DE Ra12- P510S-C construct cDNA sequence SEQ ID NO 822.
KW Human; cancer; prostate cancer; vaccine; cytostatic; immunostimulant;
KW gene therapy; gene; ss.
OS Chimeric - Mycobacterium tuberculosis.
OS Chimeric - Homo sapiens.
XX
PN US2002022248-A1.
XX
PD 21-FEB-2002.
XX
PF 12-JAN-2001; 2001US-0759143.
XX
PR 25-FEB-1997; 97US-0806099.
PR 01-AUG-1997; 97US-0904804.
PR 09-FEB-1998; 98US-0020956.
PR 25-FEB-1998; 98US-0030607.
PR 14-JUL-1998; 98US-0115453.
PR 23-SEP-1998; 98US-0159812.
PR 15-JAN-1999; 99US-0232149.
PR 09-APR-1999; 99US-0288946.
PR 13-JUL-1999; 99US-0352616.
PR 18-NOV-1999; 99US-0439313.
PR 14-JAN-2000; 2000US-0483672.
PR 27-MAR-2000; 2000US-0536857.
PR 09-MAY-2000; 2000US-0568100.
PR 12-MAY-2000; 2000US-0570737.
PR 13-JUN-2000; 2000US-0593793.
PR 27-JUN-2000; 2000US-0605783.
PR 10-AUG-2000; 2000US-0636215.
PR 29-AUG-2000; 2000US-0651236.
PR 06-SEP-2000; 2000US-0657279.
PR 02-OCT-2000; 2000US-0679426.
```

PR 10-OCT-2000; 2000US-0685166.

```
XX XUJ/J/ XU J.
PA DILL/ DILLON D C.
PA MITC/ MITCHAM J L.
PA HARL/ HARLOCKER S L.
PA JIAN/ JIANG Y.
PA KALO/ KALOS M D.
PA FANG/ FANGER G R.
PA RETT/ RETTER M W.
PA STOL/ STOLK J A.
PA DAYC/ DAY C H.
PA VEDV/ VEDVICK T S.
PA CART/ CARTER D.
PA LISX/ LI S X.
PA WANG/ WANG A.
PA SKEI/ SKEIKY Y A W.
PA HEPL/ HEPLER W T.
PA HEND/ HENDERSON R A.
XX
PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;
PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;
PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;
XX
DR WPI; 2002-255649/30.
XX
XX New prostate-specific polynucleotides for diagnosing and treating
PT diseases, in particular prostate cancer, and as markers for the
PT progression of cancer -
XX
PS Example 17; SEQ ID NO 822; 87pp; English.
XX
CC The present invention provides prostate-specific coding sequences and
CC their encoded proteins. These can be used in the diagnosis and treatment
CC of cancers, particularly prostate cancer. The present sequence is a cDNA
CC described in the invention.
XX
SQ Sequence 675 BP; 162 A; 197 C; 190 G; 126 T; 0 other;
```

Alignment Scores:

Pred. No.: 5,71e-58 Length: 675  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 24 Gaps: 0

US-09-684-215A-18 (1-128) x ABL95503 (1-675)

```
Qy 1 ThrAlaAlaSerAspAsnPhelGlnLeuSerGlnGlyGlnGlyPheAlaIleProile 20
Db 22 ACGCCGCGTCCGATAACTTCCAGCTGTCCAGGGTGGCAGGATTCCGATTCGGATC 81
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
Db 82 GGCAGGCGATGGCGATCGCGGCCAGATCAAGCTTCCACCGTTCATATCGGCGCTACC 141
Qy 41 AlaPheLeuGlyLeuGlyValValAlaAspAsnGlyAsnGlyAlaArgValGlnArgVal 60
Db 142 GCCTTCTCGGCTGGGTGGTGTTCGACAAACACGCGCAGCGGATCCACGCGGTG 201
Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
Db 202 GTCGGAGCGCTCGCGCGCAAGTCTCGCATCTCCACCGGCGAGTGCACGCGGTC 261
Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100
Db 262 GACGGCGCTCGATCAACTCGGCCACCGCATCGGCGGCGGCGCTTAACGGGCATCATCCC 321
Qy 101 GlyAspValIleSerValThrTpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
Db 322 GGTGACGTCACTCTCGGTGACCTGGCAACCAAGTCCGGCGGCGACGCGTACAGGGAACGTG 381
```

QY 121 ThrLeuAlaGluGlyProProAla 128  
 Db 382 ACATTGGCCGAGGACCCCGGCC 405

RESULT 4

ACA59940  
 ID ACA59940 standard; cDNA; 675 BP.  
 AC ACA59940;

XX 10-JUN-2003 (first entry)

DE Prostate cancer therapy associated cDNA #647.

XX Prostate cancer; vaccine; gene therapy; cytostatic; fusion protein;  
 KW immunogen; cancer; prostate specific antigen; PSA;  
 KW prostatic acid phosphatase; PAP; prostate specific membrane antigen;  
 KW PSMA; gene; ss.

OS Homo sapiens.

XX US2002192763-A1.

PN 19-DEC-2002.

XX 29-JUN-2001; 2001US-0895793.

PR 17-APR-2000; 2000US-157455P.

PR 04-OCT-2000; 2000US-0679272.

PR 28-MAR-2001; 2001US-0822827.

XX (XUJJ/) XU J.

PA (DILL/) DILLON D C.

PA (MITC/) MITCHAM J L.

PA (HARL/) HARLOCKER S L.

PA (JIAN/) JIANG Y.

PA (KALO/) KALOS M D.

PA (FANG/) FANGER G R.

PA (RETT/) RETTER M W.

PA (STOL/) STOLK J A.

PA (DAYC/) DAY C H.

PA (VEDV/) VEDVICK T S.

PA (CART/) CARTER D.

PA (LISA/) LI S X.

PA (WANG/) WANG A.

PA (SKEI/) SKEIKY Y A W.

PA (HEPL/) HEPLER W T.

PA (HEND/) HENDERSON R A.

PA (HURA/) HURAL J.

PA (MCNE/) MCNEILL P D.

PA (HOUG/) HOUGHTON R L.

PA (DBAS/) Y DE BASSOLS C V.

PA (FOVT/) FOY T M.

XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;

PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;  
 PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA, Hural J;  
 PI McNeill PD, Houghton RL, Y De Bassols CV, Foy TW;

XX WPI; 2003-352711/33.

XX New fusion protein comprising prostate-specific polypeptides, or its  
 PT immunogenic portions, useful for diagnosing, preventing and/or treating  
 PT cancer, particularly prostate cancer -

PS Example 17; SEQ ID NO 822; 85pp; English.

XX The invention describes a fusion protein comprising at least one amino  
 CC acid sequence of immunogenic portions of any of the 3 sequences not  
 CC defined in the specification, or sequences having at least 70 or 90 %  
 CC sequence identity to any one of the 35 sequences defined in the USPTO  
 CC web site, which is encoded by any of the 4 nucleotide sequences not  
 CC defined in the specification. The fusion protein, composition and

CC methods are useful for diagnosing, preventing and/or treating cancer,  
 CC particularly prostate cancer. The proteins are useful as markers to  
 CC indicate the presence or absence of cancer. This sequence  
 CC represents a prostate cancer therapy associated cDNA.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from the US patent office at  
 CC seqdata.uspto.gov/sequence.html?DocID=US20020192763.

XX Sequence 675 BP; 162 A; 197 C; 190 G; 126 T; 0 other;

Alignment Scores:

Pred. No.: 5,71e-58 Length: 675  
 Score: 653.00 Matches: 128  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 25 Gaps: 0

US-09-684-215A-18 (1-128) x ACA59940 (1-675)

QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
 Db 22 ACAGCCGCGTCCGATAACTTCCAGCTGTCACAGGTGGCAGGGATTCCCATTCGGATC 81  
 QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
 Db 82 GGGCAGGCGATGGCGATCGCGGCCAGATCAAGCTTCCACCGCTCATATCGGCGCTACC 141  
 QY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60  
 Db 142 GCCTTCCTCGCTTGGGTGTTCTGACACAAACGACGCGGACGAGTCCAACGCGTG 201  
 QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
 Db 202 GTCGGGAGCGGTCCGGGGCAAGTCTCGGCATCTCCACGGGAGGTGATCACCGGGTC 261  
 QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100  
 Db 262 GACGGCGTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAACGGGCATCATCC 321  
 QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
 Db 322 GGTGACGTCACTCGGTGACCTGGCAACCAAGTCCGGCGGCGACGCGTACAGGGAACGTC 381  
 QY 121 ThrLeuAlaGluGlyProProAla 128  
 Db 382 ACATTGGCCGAGGACCCCGGCC 405

RESULT 5

ABK39769  
 ID ABK39769 standard; cDNA; 822 BP.

AC ABK39769;

DT 21-MAY-2002 (first entry)

XX DNA encoding lung tumour protein P801P ORF5 and Ra12 fusion protein.

XX Lung tumour; cancer; T cell; immune response stimulator;

XX cytostatic; gene; ss.

OS Homo sapiens.

OS Synthetic.

PN WO200204514-A2.

XX 17-JAN-2002.

XX 10-JUL-2001; 2001WO-US22058.

PR 11-JUL-2000; 2000US-0614124.

PR 29-AUG-2000; 2000US-0651563.

PR 08-SEP-2000; 2000US-0658824.  
PR 26-SEP-2000; 2000US-0671325.  
PR 06-OCT-2000; 2000US-0677419.  
PR 30-OCT-2000; 2000US-0702705.  
PR 13-DEC-2000; 2000US-0736457.  
PR 03-MAY-2001; 2001US-0849626.  
XX (CORI-) CORIXA CORP.  
XX Wang T, Watanabe Y, Henderson RA, Johnson JC, Retter MW;  
PI Marnerakis M, Carter D, Fanger GR, Vedwick TS, Bangur CS;  
PI McNabb A, Wang A, Fanger N, Switzer A, McNeill PD, Clapper JD;  
XX WPI: 2002-164634/21.  
DR P-PSDB; AAU85587.  
XX Novel polynucleotide encoding a lung tumour polypeptide useful for  
PT stimulating and/or expanding T cells specific for a tumour protein  
XX  
PS Example 8; SEQ ID No 1862; 223pp; English.  
XX The invention describes an isolated polynucleotide and polypeptide  
CC useful for stimulating and/or expanding T cells specific for a tumour  
CC protein for determining the presence of a cancer in a patient. A  
CC composition containing the polynucleotide and/or polypeptide is useful  
CC for treating a lung cancer in a patient. The polypeptide is useful for  
CC removing tumour cells from a biological sample. The polynucleotide is  
CC also useful as probe or primer to detect the level of mRNA encoding a  
CC tumour protein. This sequence encodes a lung tumour associated protein  
CC or protein fragment, described in the method of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 822 BP; 146 A; 281 C; 270 G; 125 T; 0 other;  
Alignment Scores:  
Pred. No.: 7,23e-58 Length: 822  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 24 Gaps: 0  
US-09-684-215A-18 (1-128) x ABK39769 (1-822)  
QY 1 ThrAlaAlaSerAspAsnPhgInLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
DB 22 ACGCGCGCTCCGATACCTTCCAGCTGTCCAGGGTGGCAGGATTCGCCATTCGCATC 81  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
DB 82 GGCAGCGGATGCGATCGCGGCCAGATCAAGCTTCCACCCTTCATATCGGCGCTACC 141  
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAlaArgValGlnArgVal 60  
DB 142 GCCTTCCTCGGCTGGGTGTGTGTCGACAAACAGCGCAACGCGCAGTCCACGCGTG 201  
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
DB 202 GTCGGAGCGCTCCGGCGGAAGTCTCGGCATCTCCACCGCGGACGTGATCACCGGCGTC 261  
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100  
DB 262 GACGCGCTCCGATCACTACCTCGCGCAGCGATGGCGACGCGCTTAACGGGCATCCTCC 321  
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
DB 322 GGTGACGTCTATCTCGGTGACCTGGCAACCAACGTCGGCGGCGGCGGTACAGGAACTG 381  
QY 121 ThrLeuAlaGluGlyProProAla 128  
|||||

Db 382 ACATTGGCCGAGGACCCCGGCC 405  
RESULT 6  
ACAL2098  
ID ACAL2098 standard; cDNA; 822 BP.  
XX  
AC ACAL2098;  
XX  
DT 06-JUN-2003 (first entry)  
XX Human lung cancer protein L801P ORF5/Ral2 fusion protein cDNA.  
DE  
XX  
KW Human; lung cancer; ss; lung tumour; cytostatic; vaccine;  
KW T cell expansion; CD4; CD8; RAL2; gene.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
PN US2002197669-A1.  
XX  
PD 26-DEC-2002.  
XX  
PF 03-MAY-2001; 2001US-0849626.  
XX  
PR 13-DEC-2000; 2000US-0736457.  
XX (BANG/) BANGUR C S.  
PA (FANG/) FANGER G R.  
PA (WANG/) WANG A.  
PA (WANG/) WANG T.  
PA (SWIT/) SWITZER A P.  
PA (MCNE/) MCNEILL P D.  
PA (CLAP/) CLAPPER J D.  
XX Bangur CS, Fanger GR, Wang A, Wang T, Switzer AP, McNeill PD;  
PI Clapper JD;  
XX  
DR WPI: 2003-352750/33.  
DR P-PSDB; ABU69562.  
XX  
PT Novel lung cancer polynucleotide encoding lung cancer protein, useful  
PT for detecting the presence of lung cancer in a patient, and in  
PT pharmaceutical compositions, e.g. vaccines, for treating lung cancer  
XX  
PS Example 8; Page -: 72pp; English.  
XX The invention relates to a polynucleotide encoding a lung tumour protein,  
CC comprising a sequence selected from any of the 14 sequences  
CC mentioned in the specification, or a sequence (S2) mentioned in  
CC specification, complement of S1, sequences consisting of at least 20  
CC contiguous residues of S1, sequences that hybridise to S1, sequences  
CC having 75%, preferably 90%, identity to S1, or degenerate variants of  
CC S1. Also included are an isolated polypeptide (comprising a sequence (S3)  
CC selected from any one of the 4 amino acid sequences mentioned in the  
CC specification, a sequence encoded by the polynucleotide, or sequences  
CC having at least 70%, preferably 90%, identity to a sequence encoded by  
CC the polynucleotide), an expression vector comprising the polynucleotide  
CC operably linked to an expression control sequence, a host cell  
CC transformed or transfected with the vector, an isolated antibody (or its  
CC antigen-binding fragment) that specifically binds to the polypeptide,  
CC detecting the presence of a cancer in a patient, a fusion protein  
CC comprising the polypeptide, an oligonucleotide that hybridises to  
CC S1 under moderately stringent conditions, stimulating and/or expanding T  
CC cells specific for a tumour protein (comprising contacting T cells with  
CC the polynucleotide, protein or antigen-presenting cells, under conditions  
CC and for a time sufficient to permit the stimulation and/or expansion of T  
CC cells) and inhibiting the development of a cancer in a patient (by  
CC incubating CD4+ and/or CD8+ T cells isolated from a patient with the  
CC polynucleotide, protein or antigen presenting cells that express the  
CC polynucleotide, such that T cells proliferate, administering to the  
CC patient an effective amount of the proliferated T cells, and thus  
CC inhibiting the development of a cancer in the patient. The  
CC polynucleotide, protein and cells are useful in a composition for

CC stimulating an immune response in a patient, and for treating a cancer in  
CC a patient (particularly lung cancer). The oligonucleotide is useful for  
CC determining the presence of a cancer in a patient. The protein and  
CC oligonucleotides are useful in pharmaceutical compositions, e.g.  
CC vaccines. The polynucleotide is also useful as a probe or primer for  
CC nucleic acid hybridisation, and in the design and preparation of  
CC ribozyme molecules for inhibiting expression of tumour polypeptides and  
CC proteins in tumour cells. An amplified portion of the polynucleotide is  
CC useful for isolating a full-length gene from a suitable library.  
CC The present sequence encodes a fusion protein of human RAL2 with  
CC the protein product of a cDNA (full length, extended or partial)  
CC isolated from a library derived from lung tumour/cancer cells.  
CC Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic  
CC format directly from the USPTO  
CC at seqdata.uspto.gov/sequence.html?DocId=20020197669.

XX  
SQ Sequence 822 BP; 146 A; 281 C; 270 G; 125 T; 0 other;

## Alignment Scores:

Pred. No.: 7.23e-58 Length: 822  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 25 Gaps: 0

US-09-684-215A-18 (1-128) x ACA12098 (1-822)

QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
DB 22 AGCGCGCGTCCGATTAACCTCCAGCTGCCAGGGTGGCAGGATTCGCCATTCGCGATC 81  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
DB 82 GGGCAGGCGATGGCGATCGCGGCCAGATCAAGCTCCACCGTTCATATCGGGCTACC 141  
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArgValGlnArgVal 60  
DB 142 GCCTTCCTCGGCTTGGTGTTCACAAACGCAACGCGCAGGAGTCCACGCGTG 201  
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
DB 202 GTGCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGAGGTATCAGCGCGTC 261  
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100  
DB 262 GACGCGCTCCGATCAACTCGCCACCGGATGGCGGAGCGCTTAACGGCGATCATCC 321  
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
DB 322 GGTGACGTCTCTCGTGCACCTGGCAACCAAGTCCGGCGGACGCGTACAGGAACTG 381  
QY 121 ThrLeuAlaGluGlyProAla 128  
DB 382 ACATTTGGCGGAGGAGCCCCCGGCC 405

## RESULT 7

ACA03284  
ID ACA03284 standard; DNA; 822 BP.  
XX  
AC ACA03284;  
XX  
XX  
XX 22-MAY-2003 (first entry)  
XX  
DE Lung cancer therapy and diagnosis associated DNA #5.

XX Lung cancer; cytostatic; vaccine; gene therapy; cancer;  
KW gene; ds.  
XX  
OS Homo sapiens.

PN US2002172952-A1.

XX 21-NOV-2002.  
PD  
XX  
PF 10-JUL-2001; 2001US-0902941.  
XX  
PR 30-JUN-1999; 99US-0346492.  
PR 15-OCT-1999; 99US-0419356.  
PR 17-DEC-1999; 99US-0468867.  
PR 30-DEC-1999; 99US-0476300.  
PR 06-MAR-2000; 2000US-0519642.  
PR 22-MAR-2000; 2000US-0533077.  
PR 10-APR-2000; 2000US-0546259.  
PR 27-APR-2000; 2000US-0560406.  
PR 05-JUN-2000; 2000US-0589184.  
PR 11-JUL-2000; 2000US-0614124.  
PR 29-AUG-2000; 2000US-0651563.  
PR 08-SEP-2000; 2000US-0658824.  
PR 26-SEP-2000; 2000US-0671325.  
PR 06-OCT-2000; 2000US-0677419.  
PR 30-OCT-2000; 2000US-0702705.  
PR 13-DEC-2000; 2000US-0736457.  
PR 03-MAY-2001; 2001US-0849626.  
XX (CORI-) CORIXA CORP.  
PA  
XX  
XX Henderson RA, Wang T, Watanabe Y, Johnson JC, Retter MW, Durham M;  
PI Carter D, Fanger GR, Vedwick TS, Bangur CS, McNabb A;  
XX  
XX WPI; 2003-328427/31;  
XX  
XX New polynucleotide, useful for preparing a composition for treating or  
PT inhibiting development of cancer, e.g. lung cancer -  
PT  
XX  
PS Example 8; SEQID NO 1862; 82pp; English.  
XX  
CC The invention describes an isolated polynucleotide comprising one of 32  
CC sequences, complement or degenerate variants of them. The polynucleotide  
CC is useful for preparing a composition e.g. a vaccine or for gene therapy,  
CC for treating or inhibiting development of cancer, e.g. lung cancer.  
CC This sequence represents a polynucleotide associated with the  
CC compositions and methods for the therapy and diagnosis of lung cancer.  
XX  
SQ Sequence 822 BP; 146 A; 281 C; 270 G; 125 T; 0 other;

## Alignment Scores:

Pred. No.: 7.23e-58 Length: 822  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 25 Gaps: 0

US-09-684-215A-18 (1-128) x ACA03284 (1-822)

QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
DB 22 AGCGCGCGTCCGATTAACCTCCAGCTGCCAGGGTGGCAGGATTCGCCATTCGCGATC 81  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
DB 82 GGGCAGGCGATGGCGATCGCGGCCAGATCAAGCTTCACCGTTCATATCGGGCTACC 141  
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArgValGlnArgVal 60  
DB 142 GCCTTCCTCGGCTTGGTGTTCACAAACGCAACGCGCAGGAGTCCACGCGTG 201  
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
DB 202 GTGCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGAGGTATCAGCGCGTC 261  
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100  
DB 262 GACGCGCTCCGATCAACTCGCCACCGGATGGCGGAGCGCTTAACGGCGATCATCC 321

Qy 101 GlyAspValIleSerValThrTrpGlnThrIleSerGlyGlyThrArgThrGlyAsnVal 120  
 Db 322 GGTGACGTCATCTCGGTGACCTGGCAACCAAGTCGGCGGCACGCGTACAGGGAACGTG 381  
 Qy 121 ThrLeuAlaGluGlyProProAla 128  
 Db 382 ACATTGGCCGAGGACCCCGGCC 405

RESULT 8

ABK27798

ID ABK27798 standard; cDNA; 894 BP.

XX AC

XX ABK27798;

XX DT 09-APR-2002 (first entry)

XX DE Human colon cancer expressed sequence tag, Seq ID no 236.

XX KW Human; colon cancer; T cell expansion; tumour; EST; gene; ss;

XX KW expressed sequence tag.

XX OS Homo sapiens.

XX PN WO200196390-A2.

XX PD 20-DEC-2001.

XX PF 08-JUN-2001; 2001WO-US18577.

XX PR 09-JUN-2000; 2000US-210821P.

XX PR 18-DEC-2000; 2000US-256571P.

XX PR 10-MAY-2001; 2001US-290240P.

XX PA (CORI-) CORIXA CORP.

XX PI Jiang Y, Hepler WT, Clapper JD, Wang A, Secrist H;

XX DR WPI; 2002-139708/18.

XX DR P-PSDB; AAU81532.

XX PT Novel isolated polynucleotide encoding a polypeptide comprising a

XX PT portion of colon tumour protein, useful for detection, diagnosis and

XX PT therapy of human colon cancer

XX PS Claim 1; Page 215-216; 220pp; English.

XX CC The invention relates to an isolated polynucleotide (I) encoding a

XX CC polypeptide (II) comprising at least a portion of a colon tumour

XX CC protein. (I), (II) and antibody (III) to (II) are useful for determining

XX CC the presence of a cancer in a patient. (I), (II) or antigen presenting

XX CC cells expressing (I) is useful for stimulating and/or expanding T cells

XX CC specific for a tumour protein, by contacting T cells with (I), (II) or

XX CC antigen-presenting cells that express (I), under conditions and for a

XX CC time sufficient to permit the stimulation and/or expansion of T cells.

XX CC (I), (II), or antigen presenting cells that express (II) are useful for

XX CC treating colon cancer in a patient by incubating CD4+ and/or CD8+ T cells

XX CC isolated from a patient with (I), (II) or antigen presenting cells that

XX CC express (II), such that T cells proliferate, and administering to the

XX CC patient an effective amount of the proliferated T cells, thus inhibiting

XX CC the development of a cancer in the patient. (I) or (II) is useful in

XX CC vaccines and pharmaceutical compositions for prevention and treatment

XX CC of colon malignancies and for the diagnosis and monitoring of such

XX CC cancers. (I), (II) or (III) is useful for detection, diagnosis and/or

XX CC therapy of human colon cancer. (I) is useful as a probe or primer for

XX CC nucleic acid hybridisation, and in the design and preparation of

XX CC ribozyme molecules for inhibiting expression of (II) in tumour cells.

XX CC ABK27364-ABK27807 represent novel human colon cancer coding

XX CC sequences and primers of the invention.

XX SQ Sequence 894 BP; 207 A; 269 C; 237 G; 181 T; 0 other;

Alignment Scores:

Pred. No.: 7.99e-58 Length: 894  
 Score: 653.00 Matches: 128  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 24 Gaps: 0  
 US-09-684-215A-18 (1-128) x ABK27798 (1-894)  
 Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20  
 Db 22 ACGCCCGCGTCCGATAACTTCCAGCTGTCCAGGTGGCAGGATTCCGCAATTCCTCCATC 81  
 Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLeuProThrValHisIleGlyProThr 40  
 Db 82 GGCAGGCGATGCGGATCGCGGCCAGATCAAGCTTCCACCGCTTCATATCGGGCCCTACC 141  
 Qy 41 AlaphLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArgValGlnArgVal 60  
 Db 142 GCCTTCTCGCTGGGTGGTGTGTCGACACACGCGCAACGCGCAGTCCACGCGGTG 201  
 Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
 Db 202 GTCGGGAGCGCTCGCGCGCAAGTCTCGGCAATCTCCACCGCGGACGTGATCACCGCGTC 261  
 Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100  
 Db 262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGCGCGACGCGCTTAACGGGCATCATCCC 321  
 Qy 101 GlyAspValIleSerValThrTrpGlnThrIleSerGlyGlyThrArgThrGlyAsnVal 120  
 Db 322 GGTGACGTCATCTCGGTGACCTGGCAACCAAGTCGGCGGCACGCGTACAGGGAACGTG 381  
 Qy 121 ThrLeuAlaGluGlyProProAla 128  
 Db 382 ACATTGGCCGAGGACCCCGGCC 405

RESULT 9

ABQ92443

ID ABQ92443 standard; DNA; 900 BP.

XX AC

XX ABQ92443;

XX DT 07-OCT-2002 (first entry)

XX DE Human lung cancer associated DNA sequence SEQ ID NO:353.

XX KW Human; lung cancer; lung tumour; cytostatic; gene therapy; vaccine;

XX KW gene; ds.

XX OS Homo sapiens.

XX PN WO200247534-A2.

XX PD 20-JUN-2002.

XX PF 30-NOV-2001; 2001WO-US47576.

XX PR 12-DEC-2000; 2000US-0735705.

XX PR 07-MAY-2001; 2001US-0850716.

XX PR 28-JUN-2001; 2001US-0897778.

XX PA (CORI-) CORIXA CORP.

XX PI Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;

XX PI McNeill PD, Fanger N, Retter MW, Durham M, Fanger GR, Vedwick TS;

XX PI Carter D, Watanabe Y, Peckham DW, Cal F, Foy TM;

XX DR WPI; 2002-583465/62.

XX DR P-PSDB; ABP61920.

XX PT Novel lung carcinoma polynucleotide sequences and polypeptides encoded

XX PT by the polynucleotides, useful in pharmaceutical compositions such as

PT vaccines and as markers to indicate the presence of lung cancer -  
 PS Claim 8; Page 340-341; 381pp; English.  
 XX  
 CC The present invention describes isolated human lung carcinoma  
 CC polynucleotides (I) and polypeptides (II). (I) and (II) have cytostatic  
 CC activity, and can be used in gene therapy and in vaccines. Compositions  
 CC comprising (I) or (II) can be used for stimulating an immune response in  
 CC a patient and for treating lung cancer in a patient. Oligonucleotides of  
 CC (I) can be used for detecting the presence of a cancer in a patient, by  
 CC obtaining a biological sample from the patient, contacting the  
 CC biological sample with the oligonucleotide, detecting in the sample, an  
 CC amount of polynucleotide that hybridises to the oligonucleotide and  
 CC comparing the amount of polynucleotide that hybridises to the  
 CC oligonucleotide to a predetermined cut-off value, and determining the  
 CC presence of a cancer in the patient. (I) and (II) are useful in  
 CC pharmaceutical compositions, e.g. vaccines. (I) is useful as a marker to  
 CC indicate the presence or absence of a cancer such as lung cancer.  
 CC ABQ92145 to ABQ92486 and ABP61866 to ABP61992 represent sequences used  
 CC in the exemplification of the present invention.  
 XX  
 SQ Sequence 900 BP; 219 A; 239 C; 246 G; 196 T; 0 other;

Alignment Scores:  
 Pred. No.: 8,06e-58 Length: 900  
 Score: 653.00 Matches: 128  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 24 Gaps: 0

US-09-684-215A-18 (1-128) x ABQ92443 (1-900)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20  
 DB 22 ACGGCGCGCTCCGATTAACCTCCAGCTGTCCAGGTTGGCAGGATTCGCCATTCGATC 81  
 QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
 DB 82 GGGCAGCGATGGCGATCGCGGCGCAGATCAAGCTTCCACCGCTTCATATCGGGCTACC 141  
 QY 41 AlaPheLeuGlyLeuGlyValValAlaAspAsnGlyAsnGlyAlaArgValGlnArgVal 60  
 DB 142 GCCTTCTCGGCTTGGTGTTCGACAAACACGCGACGCGAGTCCCAACGCGTG 201  
 QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
 DB 202 GTCCGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACGCGGACGTGATCACCAGCGGTC 261  
 QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100  
 DB 262 GACGCGCTCCGATCAACTCGCGCACCGCGATGGCGGACGCGCTTAACGGGCATCATCC 321  
 QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyThrArgThrGlyAsnVal 120  
 DB 322 GTGAGCTCATCTCGGTGACCTGGCAACCAACCAAGTCCGGCGGCGACGCGTACAGGAGCGTG 381  
 QY 121 ThrLeuAlaGluGlyProProAla 128  
 DB 382 ACATTGGCGAGGAGCCCCCGGCC 405

RESULT 10

ABL49257  
 ID ABL49257 standard; cDNA; 900 BP.

XX  
 AC ABL49257;

XX  
 DT 01-MAY-2002 (first entry)

DE Ral2/C-terminal portion of L763p fusion polynucleotide SEQ ID NO:353.  
 XX Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;  
 KW immune response; ss.

XX Homo sapiens.  
 OS WO200200174-A2.  
 PN 03-JAN-2002.  
 XX 28-JUN-2001; 2001WO-US21065.  
 XX 28-JUN-2000; 2000US-0606421.  
 PR 02-AUG-2000; 2000US-0630940.  
 PR 21-AUG-2000; 2000US-0643597.  
 PR 15-SEP-2000; 2000US-0662786.  
 PR 09-OCT-2000; 2000US-0685696.  
 PR 12-DEC-2000; 2000US-0735705.  
 PR 07-MAY-2001; 2001US-0850716.  
 XX (CORI-) CORIXA CORP.  
 PA Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;  
 PI McNeill PD, Fanger-N, Retter MW, Marnerakis M, Fanger GR;  
 PI Vedvick TS, Carter D, Watanabe Y, Peckham DW;  
 XX WPI: 2002-090513/12.  
 DR P-PSDB; ABB75000.  
 XX Polynucleotides encoding lung tumor polypeptides, useful for treating  
 PT lung cancer or stimulating an immune response -  
 PT Claim 1; Page 333-334; 374pp; English.  
 PS The present invention describes human lung tumour proteins. Human lung  
 CC tumour proteins and polynucleotides have cytostatic and immunostimulant  
 CC activities, and can be used in vaccine production. Compositions  
 CC comprising the lung tumour proteins, polynucleotides, antibodies,  
 CC fusion proteins, T cell populations, or antigen presenting cells that  
 CC express the lung tumour proteins are useful for treating lung cancer or  
 CC stimulating an immune response. ABL49959 to ABL49300 and ABB74946 to  
 CC ABB75070 represent sequences used in the exemplification of the present  
 CC invention.  
 XX SQ Sequence 900 BP; 219 A; 239 C; 246 G; 196 T; 0 other;

Alignment Scores:  
 Pred. No.: 8,06e-58 Length: 900  
 Score: 653.00 Matches: 128  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 24 Gaps: 0

US-09-684-215A-18 (1-128) x ABL49257 (1-900)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20  
 DB 22 ACGGCGCGCTCCGATTAACCTCCAGCTGTCCAGGTTGGCAGGATTCGCCATTCGATC 81  
 QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
 DB 82 GGGCAGCGATGGCGATCGCGGCGCAGATCAAGCTTCCACCGCTTCATATCGGGCTACC 141  
 QY 41 AlaPheLeuGlyLeuGlyValValAlaAspAsnGlyAsnGlyAlaArgValGlnArgVal 60  
 DB 142 GCCTTCTCGGCTTGGTGTTCGACAAACACGCGACGCGAGTCCCAACGCGTG 201  
 QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
 DB 202 GTCCGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACGCGGACGCGTACAGGAGCGTG 261  
 QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100  
 DB 262 GACGCGCTCCGATCAACTCGCGCACCGCGATGGCGGACGCGCTTAACGGGCATCATCC 321

Qy 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
 Db 322 GGTGACGTCTCTCGGTGACCTGGCAACCAAGTCGGCGCGACGCTACAGGGAACGTG 381  
 Qy 121 ThrLeuAlaGluGlyProProAla 128  
 Db 382 ACATTGGCCGAGGAGCCCGGCC 405

RESULT 11  
 ID AAS64141 standard; cDNA; 915 BP.  
 XX  
 AC AAS64141;  
 XX  
 XX 29-JAN-2002 (first entry)  
 XX  
 XX Human /M. tuberculosis Ra12 fusion protein RA12-P775P-ORF3 cDNA.  
 XX  
 XX Human; prostate cancer; ss; cytostatic; immunostimulant; tumour.  
 XX  
 OS Chimeric - Homo sapiens.  
 OS Chimeric - Microbacterium tuberculosis.  
 OS Synthetic.  
 OS  
 PN WO200173032-A2.  
 XX  
 PD 04-OCT-2001.  
 XX  
 XX 27-MAR-2001; 2001WO-US09919.  
 XX  
 XX 27-MAR-2000; 2000US-0536857.  
 XX  
 PR 09-MAY-2000; 2000US-0568100.  
 PR  
 PR 12-MAY-2000; 2000US-0570737.  
 PR  
 PR 13-JUN-2000; 2000US-0593793.  
 PR  
 PR 27-JUN-2000; 2000US-0605783.  
 PR  
 PR 10-AUG-2000; 2000US-0636215.  
 PR  
 PR 29-AUG-2000; 2000US-0851236.  
 PR  
 PR 06-SEP-2000; 2000US-0857279.  
 PR  
 PR 02-OCT-2000; 2000US-0679426.  
 PR  
 PR 10-OCT-2000; 2000US-0685166.  
 XX  
 XX (CORI-) CORIXA CORP.  
 XX  
 XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;  
 PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;  
 PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;  
 XX  
 XX WPI: 2001-639232/73.  
 DR  
 DR P-PSDB; AAU69902.  
 XX  
 XX New human prostate-specific polypeptides and polynucleotides useful for  
 PT the diagnosis and treatment of cancer, especially prostate cancer -  
 XX  
 XX Example 17; Page 536-537; 579pp; English.  
 XX  
 CC The invention relates to isolated prostate-specific  
 CC polynucleotides, polypeptides, fusion proteins of the polypeptides,  
 CC antibodies raised against the polypeptides (or antigenic epitopes  
 CC derived from them) and antigen-presenting cells expressing the  
 CC polypeptides. The antibodies are useful for detecting the presence of  
 CC cancer, especially prostate cancer. The polypeptides, polynucleotides and  
 CC the antigen-presenting cells are useful for stimulating and/or expanding  
 CC T cells specific for a tumour protein, and for inhibiting the development  
 CC of cancer especially prostate cancer. Compositions comprising the  
 CC polynucleotide and/or polypeptide are useful for stimulating an immune  
 CC response, and for treating cancer. The oligonucleotide is useful for  
 CC detecting cancer. The present sequence is a prostate specific  
 CC polynucleotide of the invention.  
 XX  
 SQ Sequence 915 BP; 167 A; 280 C; 284 G; 184 T; 0 other;

Alignment Scores:  
 Pred. No.: 8.22e-58 Length: 915

Score: 653.00 Matches: 128  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 Db: 22 Gaps: 0

US-09-684-215A-18 (1-128) x AAS64141 (1-915)

Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20  
 Db 22 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGGTGGGAGGATTCCGATCCGATC 81  
 Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
 Db 82 GGGCAGCGCATGGCGATCGCGGCGCAGATCAAGCTTCCACCGTTCATATCGGCGCTACC 141  
 Qy 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60  
 Db 142 GCCTTCCTCGGCTTGGGTGTTCGACAAACGCAACGCGCAGGTCACGCGGCGT 201  
 Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
 Db 202 GTCGGAGCGCTCCGCGCGCAAGTCTCGGCATCTCCACCGCGGAGCGTATCACGCGGTC 261  
 Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100  
 Db 262 GACGGCGCTCCGATCAACTCGGCGCACCGCGATGCGGCGCGCTTAACGGGCGATCATCCC 321  
 Qy 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
 Db 322 GGTGACGTCTCTCGGTGACCTGGCAACCAAGTCGGCGCGACGCTACAGGGAACGTG 381

RESULT 12  
 AAH93905  
 ID AAH93905 standard; cDNA; 915 BP.  
 XX  
 AC AAH93905;  
 XX  
 DT 04-OCT-2001 (first entry)  
 XX  
 DE Ra12-P775P-ORF3 construct cDNA sequence.  
 XX  
 KW Human; prostate cancer; prostate-specific; diagnosis; vaccine;  
 KW cytostatic; gene therapy; metastasis; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200151633-A2.  
 XX  
 PD 19-JUL-2001.  
 XX  
 PF 16-JAN-2001; 2001WO-US01574.  
 XX  
 PR 14-JAN-2000; 2000US-0483672.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;  
 PI Kalos MD, Fanger GR, Day CH, Retter MW, Stolk JA, Skeiky YAW;  
 PI Wang A, Mesagher MJ;  
 XX  
 DR WPI: 2001-425873/45.  
 XX  
 XX New polynucleotide encoding a prostate-specific protein, for  
 PT diagnosing, monitoring and treating prostate cancer in a patient and  
 PT for use in vaccines -  
 XX  
 PS Claim 8; Page 497-498; 543pp; English.  
 XX

CC The present invention describes polynucleotide sequences (I) which encode  
 CC prostate-specific proteins (II). (I) and (II) have cytostatic activity,  
 CC and can be used in vaccine production and gene therapy. (I), (II),  
 CC antibodies to (II), fusion proteins comprising (II), and isolated  
 CC T cells prepared using (I) or (II) are used treat cancer in a patient.  
 CC (I) and the antibodies are also used in the detection of cancer in a  
 CC patient. The cancer that is diagnosed or treated is particularly  
 CC prostate cancer. (I) and (II) can be used in vaccines. The antibodies or  
 CC (I) can be used for monitoring the progression of cancer in a patient.  
 CC (I) and (II) can also be used to improve diagnostic and therapeutic  
 CC methods for prostate cancer. They can indicate the level of metastasis  
 CC as well as the prostate volume. AH93357 to AAH93944 and AAH01115 to  
 CC AAH01318 represent polynucleotide and amino acid sequences used in the  
 CC exemplification of the present invention.

XX Sequence 915 BP; 167 A; 280 C; 284 G; 184 T; 0 other;

## Alignment Scores:

Pred. No.: 8,22e-58 Length: 915  
 Score: 653.00 Matches: 128  
 Percent Similarity: 100.00% Conservativeness: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 22 Gaps: 0

US-09-684-215A-18 (1-128) x AAH93905 (1-915)

QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20  
 DB 22 ACGGGCGGCTCCGATAACTCCAGCTGTCCAGGGTGGCAGGGATTCGCCATTCGGATC 81  
 QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
 DB 82 GGGCAGGGCATCGCATCGGGCCAGATCAAGCTCCACCGCTTCATATCGGGCTACC 141  
 QY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60  
 DB 142 GCCTTCCTCGGCTTGGGTGTGTCACACACGACGAGCGGCAGGATCCCAACGGCGTG 201  
 QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
 DB 202 GTCCGGAGGCGTCCGGCGCAAGTCTCGGCATCTCCACGGCGAGTGATCACCGCGTC 261  
 QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100  
 DB 262 GACGGCGTCCGATCAACTCGGCCACCGGATGGCGGACGCGCTTAACGGGCATCATCC 321  
 QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
 DB 322 GGTGACGTCATCTCGGTGACCTGGCAACCAAGTCCGGCGGCACCGGTACAGGGACGTCG 381  
 QY 121 ThrLeuAlaGluGlyProProAla 128  
 DB 382 ACATTGGCGGAGGACCCCGGCC 405

## RESULT 13

ABL95512  
 ID ABL95512 standard; cDNA; 915 BP.

AC ABL95512;

DT 19-JUL-2002 (first entry)

DE Ra12- P775P-ORF3 construct cDNA sequence SEQ ID NO 834.

XX Human; cancer; prostate cancer; vaccine; cytostatic; immunostimulant;  
 KW gene therapy; gene; ss.

XX Chimeric - Mycobacterium tuberculosis.

OS Chimeric - Homo sapiens.

XX US2002022248-A1.

XX

PD 21-FEB-2002.  
 XX 12-JAN-2001; 2001US-0759143.  
 PF 25-FEB-1997; 97US-0806099.  
 XX 01-AUG-1997; 97US-0904804.  
 PR 09-FEB-1998; 98US-0020956.  
 PR 25-FEB-1998; 98US-0030607.  
 PR 14-JUL-1998; 98US-0115453.  
 PR 23-SEP-1998; 98US-0159812.  
 PR 15-JAN-1999; 99US-0232149.  
 PR 09-APR-1999; 99US-0288946.  
 PR 13-JUL-1999; 99US-0352616.  
 PR 12-NOV-1999; 99US-0439313.  
 PR 18-NOV-1999; 99US-0443686.  
 PR 14-JAN-2000; 2000US-0483672.  
 PR 27-MAR-2000; 2000US-0536857.  
 PR 09-MAY-2000; 2000US-0568100.  
 PR 12-MAY-2000; 2000US-0570737.  
 PR 13-JUN-2000; 2000US-0593793.  
 PR 27-JUN-2000; 2000US-0605783.  
 PR 10-AUG-2000; 2000US-0636215.  
 PR 29-AUG-2000; 2000US-0651236.  
 PR 06-SEP-2000; 2000US-0657279.  
 PR 02-OCT-2000; 2000US-0679426.  
 PR 10-OCT-2000; 2000US-0685166.  
 XX (XUJJ/) XU J.  
 PA (DILL/) DILLON D C.  
 PA (MITC/) MITCHAM J L.  
 PA (HARL/) HARLOCKER S L.  
 PA (JIAN/) JIANG Y.  
 PA (KALO/) KALOS M D.  
 PA (FANG/) FANGER G R.  
 PA (RETT/) RETTER M W.  
 PA (STOL/) STOLK J A.  
 PA (DAYC/) DAY C H.  
 PA (VEDV/) VEDVICK T S.  
 PA (CART/) CARTER D.  
 PA (LISX/) LI S X.  
 PA (WANG/) WANG A.  
 PA (SKEI/) SKEIKY Y A W.  
 PA (HEPL/) HEPLER W T.  
 PA (HEND/) HENDERSON R A.  
 XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;  
 PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;  
 PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;  
 XX WPI; 2002-255649/30.

XX New prostate-specific polynucleotides for diagnosing and treating  
 PT diseases, in particular prostate cancer, and as markers for the  
 PT progression of cancer

PS Example 17; SEQ ID NO 834; 87pp; English.

XX The present invention provides prostate-specific coding sequences and  
 CC their encoded proteins. These can be used in the diagnosis and treatment  
 CC of cancers, particularly prostate cancer. The present sequence is a cDNA  
 CC described in the invention.

XX Sequence 915 BP; 167 A; 280 C; 284 G; 184 T; 0 other;

## Alignment Scores:

Pred. No.: 8,22e-58 Length: 915  
 Score: 653.00 Matches: 128  
 Percent Similarity: 100.00% Conservativeness: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 24 Gaps: 0

US-09-684-215A-18 (1-128) x ABL95512 (1-915)



Qy 1 ThrAlaAlaSerAspAsnPhenGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
 Db 22 ACGGCGCGTCCGATACCTCCAGCTGTCCAGGGTGGCAGGATTCGCATTCGGATC 81  
 Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
 Db 82 GGGCAGCGATGCGATGCGGCGCCAGATCAAGCTTCCACCGCTTCATATCGGCGCTACC 141  
 Qy 41 AlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaAraGValGlnArgVal 60  
 Db 142 GCCTTCCTCGGCTTGGGTGTTCGACAAACACGCGGCGGCGGACGATCCACGCGGTG 201  
 Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
 Db 202 GTCGGGAGCGTCCGCGCGCAAGTCTCGCATCTCCACCGCGGCGATCACCGCGGTC 261  
 Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100  
 Db 262 GACGGCGCTCCGATCAACTCCGCGCACCGGATGCGGCGGCGCTTAACGGGCGATCAATCC 321  
 Qy 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
 Db 322 GGTGACGTCTATCTCGGTGACCTGGCAACCAAGTCGGCGGCGGCGGCGGTACAGGGAAGCTG 381  
 Qy 121 ThrLeuAlaGluGlyProProAla 128  
 Db 382 ACATTGGCGGAGGACCCCGGCGC 405

RESULT 14  
 ACA59949  
 ID ACA59949 standard; cdna; 915 BP.  
 XX AC ACA59949;  
 DT 10-JUN-2003 (first entry)  
 XX Prostate cancer therapy associated cdna #650.  
 DE Prostate cancer; vaccine; gene therapy; cytostatic; fusion protein;  
 KW immunogen; cancer; prostate specific antigen; PSA;  
 KW prostatic acid phosphatase; PAP; prostate specific membrane antigen;  
 KW PSMA; gene; ss.  
 XX Homo sapiens.  
 OS  
 XX US2002192763-A1.  
 PN  
 XX 19-DEC-2002.  
 PD  
 XX 29-JUN-2001; 2001US-0895793.  
 XX 17-APR-2000; 2000US-157455P.  
 PR 04-OCT-2000; 2000US-0679272.  
 PR 28-MAR-2001; 2001US-0822827.  
 XX (XUJ)/ XU J.  
 PA (DILL)/ DILLON D C.  
 PA (MITC)/ MITCHAM J L.  
 PA (HARL)/ HARLOCKER S L.  
 PA (JIAN)/ JIANG Y.  
 PA (KALO)/ KALOS M D.  
 PA (FANG)/ FANGER G R.  
 PA (RETT)/ RETTER M W.  
 PA (STOL)/ STOLK J A.  
 PA (DAYC)/ DAY C H.  
 PA (VEDV)/ VEDVICK T S.  
 PA (CART)/ CARTER D.  
 PA (LISX)/ LI S X.  
 PA (WANG)/ WANG A.  
 PA (SKEI)/ SKEIKY Y A W.  
 PA (HEPL)/ HEPLER W T.  
 PA (HEND)/ HENDERSON R A.

PA (HURA)/ HURAL J.  
 PA (MCNE)/ MCNEILL P D.  
 PA (HOUG)/ HOUGHTON R L.  
 PA (DBAS)/ Y DE BASSOLS C V.  
 PA (FOYT)/ FOY T M.  
 XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;  
 PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;  
 PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA, Hural J;  
 PI Mcneill PD, Houghton RL, Y De Bassols CV, Foy TM;  
 XX WPI: 2003-352711/33.  
 XX New fusion protein comprising prostate-specific polypeptides, or its  
 PT immunogenic portions, useful for diagnosing, preventing and/or treating  
 PT cancer, particularly prostate cancer  
 XX Example 17; SEQ ID NO 834; 85pp; English.  
 PS The invention describes a fusion protein comprising at least one amino  
 CC acid sequence of immunogenic portions of any of the 3 sequences not  
 CC defined in the specification, or sequences having at least 70 or 90 %  
 CC sequence identity to any one of the 35 sequences defined in the USPTO  
 CC web site, which is encoded by any of the 4 nucleotide sequences not  
 CC defined in the specification. The fusion protein, composition and  
 CC methods are useful for diagnosing, preventing and/or treating cancer,  
 CC particularly prostate cancer. The proteins are useful as markers to  
 CC indicate the presence or absence of cancer. This sequence  
 CC represents a prostate cancer therapy associated cdna.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from the US patent office at  
 CC seqdata.uspto.gov/sequence.html?DocID=US20020192763.  
 XX SQ Sequence 915 BP; 167 A; 280 C; 284 G; 184 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 8.22e-58 Length: 915  
 Score: 653.00 Matches: 128  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 25 Gaps: 0  
 US-09-684-215A-18 (1-128) x ACA59949 (1-915)  
 Qy 1 ThrAlaAlaSerAspAsnPhenGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
 Db 22 ACGGCGCGTCCGATACCTCCAGCTGTCCAGGGTGGCAGGATTCGCATTCGGATC 81  
 Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
 Db 82 GGGCAGCGATGCGATGCGGCGCCAGATCAAGCTTCCACCGCTTCATATCGGCGCTACC 141  
 Qy 41 AlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaAraGValGlnArgVal 60  
 Db 142 GCCTTCCTCGGCTTGGGTGTTCGACAAACACGCGGCGGCGGACGATCCACGCGGTG 201  
 Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
 Db 202 GTCGGGAGCGTCCGCGCGCAAGTCTCGCATCTCCACCGCGGCGGATCACCGCGGTC 261  
 Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100  
 Db 262 GACGGCGCTCCGATCAACTCCGCGCACCGGATGCGGCGGCGCTTAACGGGCGATCAATCC 321  
 Qy 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
 Db 322 GGTGACGTCTATCTCGGTGACCTGGCAACCAAGTCGGCGGCGGCGGCGGTACAGGGAAGCTG 381  
 Qy 121 ThrLeuAlaGluGlyProProAla 128  
 Db 382 ACATTGGCGGAGGACCCCGGCGC 405

## RESULT 15

ABK39768  
ID ABK39768 standard; cDNA; 945 BP.XX AC  
XX ABK39768;

XX DT 21-MAY-2002 (first entry)

XX DE DNA encoding lung tumour protein p801p ORF4 and Ra12 fusion protein.

XX KW Lung tumour; cancer; T cell; immune response stimulator;  
KW Cytostatic; gene; ss.XX OS Homo sapiens.  
OS Synthetic.

XX PN WO200204514-A2.

XX PD 17-JAN-2002.

XX PF 10-JUL-2001; 2001WO-US22058.

XX PR 11-JUL-2000; 2000US-0614124.

XX PR 29-AUG-2000; 2000US-0651563.

XX PR 08-SEP-2000; 2000US-0658824.

XX PR 26-SEP-2000; 2000US-0671325.

XX PR 06-OCT-2000; 2000US-0677419.

XX PR 30-OCT-2000; 2000US-0703705.

XX PR 13-DEC-2000; 2000US-0736457.

XX PR 03-MAY-2001; 2001US-0849626.

(CORI-) CORIXA CORP.

XX Wang T, Watanabe Y, Henderson RA, Johnson JC, Retter MW;

PI Marnerakis M, Carter D, Fanger CR, Vedvick TS, Bangur CS;

PI Mcnabb A, Wang A, Fanger N, Switzer A, Mcneill PD, Clapper JD;

XX WPI: 2002-164634/21.

DR P-PSDB; AAU85586.

XX Novel polynucleotide encoding a lung tumour polypeptide useful for  
PT stimulating and/or expanding T cells specific for a tumour protein

XX Example 8; SEQ ID No 1861; 223pp; English.

XX The invention describes an isolated polynucleotide and polypeptide  
CC useful for stimulating and/or expanding T cells specific for a tumour  
CC protein for determining the presence of a cancer in a patient. A  
CC composition containing the polynucleotide and/or polypeptide is useful  
CC for treating a lung cancer in a patient. The polypeptide is useful for  
CC removing tumour cells from a biological sample. The polynucleotide is  
CC also useful as probe or primer to detect the level of mRNA encoding a  
CC tumour protein. This sequence encodes a lung tumour associated protein  
CC or protein fragment, described in the method of the invention.CC Note: the sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 945 BP; 150 A; 318 C; 360 G; 117 T; 0 other;

## Alignment Scores:

Pred. No.:	8.54e-58	Length:	945
Score:	653.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	24	Gaps:	0

US-09-684-215A-18 (1-128) x ABK39768 (1-945)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20

Db	22	ACGGCCGGTCCGATAACTTCCAGGTGTCAGAGGTGGCAGGGATTCCGCCATTCCGATC	81
Qy	21	GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr	40
Db	82	GGGCAGGGGATGGCGATCGCGGGCCAGATCAAGCTTCCACCGCTTCATATCGGGCTACC	141
Qy	41	AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal	60
Db	142	GCCTTCTCTGGCTTGGGTGTTCTGCACAACAACGACGCGGCGCAGAGTCCCAACGCGTG	201
Qy	61	ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal	80
Db	202	GTCCGGAGCGTCCCGCGCAAGTCTCGGCATCTCCACGGGACGTGATCACCGCGGTC	261
Qy	81	AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro	100
Db	262	GACGGCGTCCGATCAACTCGGCCACCGCGGATGGCGGACGCCCTTAACGGGCGATCATCC	321
Qy	101	GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal	120
Db	322	GGTGACGTCACTCTCGGTGACCTGGCAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG	381
Qy	121	ThrLeuAlaGlyGlyProAla	128
Db	382	ACATTGGCGCGAGGGACCCCGCGCC	405

Search completed: September 5, 2003, 09:04:19  
Job time : 250.83 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 5, 2003, 08:28:59 ; Search time 68.4138 Seconds  
(without alignments)  
825.814 Million cell updates/sec

Title: US-09-684-215A-18

Perfect score: 653

Sequence: 1 TAASDNFQLSQGGQGAIP1.....QTKSGGTRTGNVTLAGPPA 128

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-O=/cgn2.1/USPTO\_spool\_p/US09684215/runat\_05092003\_072203\_593/app\_query.fasta\_1.853  
-DB-Issued\_Patents\_NA -QWTF-fastcap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bloms62 -TRANS=human40.cdi  
-LIST=45 -LOCAL=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFTW=ptc -NORM-ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09684215\_8CGN\_1\_145 @runat\_05092003\_072203\_593 -NCPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued\_Patents\_NA:\*  
1: /cgn2.6/ptodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2.6/ptodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2.6/ptodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2.6/ptodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2.6/ptodata/1/ina/PCTUS\_COMB.seq:\*  
6: /cgn2.6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	653	100.0	822	4	US-09-736-457-1862
2	653	100.0	900	4	Sequence 1862, App
3	653	100.0	900	4	Sequence 353, App
4	653	100.0	945	4	US-09-606-421B-353
5	653	100.0	1012	4	Sequence 1861, App
6	653	100.0	1012	4	US-09-643-597-351
7	653	100.0	1464	4	Sequence 351, App
8	653	100.0	1464	4	US-09-606-421B-351
9	653	100.0	1464	4	US-09-598-419-348
10	653	100.0	1557	4	Sequence 348, App
11	653	100.0	1557	4	US-09-620-412C-332
12	653	100.0	1578	4	Sequence 332, App
					Sequence 188, App

13	653	100.0	1578	4	US-09-598-419-188
14	653	100.0	1752	4	US-09-620-412C-352
15	653	100.0	1752	4	Sequence 352, App
16	653	100.0	1758	4	US-09-598-419-352
17	653	100.0	1758	4	Sequence 336, App
18	653	100.0	1860	4	US-09-620-412C-336
19	653	100.0	1860	4	US-09-598-419-308
20	653	100.0	1896	4	US-09-620-412C-308
21	653	100.0	1896	4	Sequence 324, App
22	653	100.0	1941	4	US-09-598-419-324
23	653	100.0	1941	4	Sequence 316, App
24	653	100.0	1965	4	US-09-620-412C-316
25	653	100.0	1965	4	Sequence 340, App
26	653	100.0	2052	4	US-09-598-419-340
27	653	100.0	2052	4	Sequence 356, App
28	653	100.0	2076	4	US-09-620-412C-356
29	653	100.0	2076	4	Sequence 312, App
30	653	100.0	2103	4	US-09-598-419-312
31	653	100.0	2103	4	Sequence 312, App
32	653	100.0	2148	4	US-09-620-412C-344
33	653	100.0	2148	4	Sequence 344, App
34	653	100.0	2148	4	US-09-598-419-344
35	653	100.0	2148	4	Sequence 320, App
36	653	100.0	2148	4	US-09-620-412C-320
37	653	100.0	2148	4	Sequence 328, App
38	653	100.0	2148	4	Sequence 320, App
39	653	100.0	2148	4	Sequence 328, App
40	653	100.0	2148	4	Sequence 328, App
41	653	100.0	2148	4	Sequence 328, App
42	653	100.0	2148	4	Sequence 328, App
43	653	100.0	2148	4	Sequence 328, App
44	653	100.0	2148	4	Sequence 328, App
45	653	100.0	2148	4	Sequence 328, App

#### ALIGNMENTS

#### RESULT 1

US-09-736-457-1862  
Sequence 1862, Application US/09736457

Patent No. 6509448

GENERAL INFORMATION:

APPLICANT: Wang, Tongtong  
APPLICANT: Bangur, Chaitanya S.  
APPLICANT: Lodes, Michael A.  
APPLICANT: Fanger, Gary  
APPLICANT: Vedvick, Tom  
APPLICANT: Carter, Darrick  
APPLICANT: Retter, Marc  
APPLICANT: Mannion, Jane  
APPLICANT: Fan, Liqun  
APPLICANT: Wang, Aijun

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER

FILE REFERENCE: 210121.478C15

CURRENT APPLICATION NUMBER: US/09/736,457

CURRENT FILING DATE: 2000-12-13

NUMBER OF SEQ ID NOS: 1864

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 1862

LENGTH: 822

TYPE: DNA

ORGANISM: Homo sapiens

US-09-736-457-1862

Alignment Scores:

Pred. No.: 1.9e-64

Score: 653.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 100.00%

DB:

Length: 822  
Matches: 128  
Conservative: 0  
Mismatches: 0  
Indels: 0  
Gaps: 0



```
Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
Db 202 GTCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGAGCGTGATCACCGCGGTC 261
Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100
Db 262 GACGGCGCTCCGATCAACTCGGCCACCGCATGGCGGACGCGCTTAACGGGCATCATCCC 321
Qy 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
Db 322 GGTGACGTCTCTCGGTGACCTGGCAACCAAGTCGGCGCGCACGCTACAGGACGCTG 381
Qy 121 ThrLeuAlaGluGlyProProAla 128
Db 382 ACATTGGCGGAGGACCCCGGCC 405

RESULT 4
US-09-736-457-1861
; Sequence 1861, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedrick, Tom
; APPLICANT: Carter, Derrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1861
; LENGTH: 945
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-736-457-1861

Alignment Scores:
Pred. No.: 2,28e-64 Length: 945
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-684-215A-18 (1-128) x US-09-736-457-1861 (1-945)

Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20
Db 22 ACGGCGCGCTCCGATAACTTCCAGCTGTCCAGGGTGGCGAGGATTGCCATTCCCGATC 81
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
Db 82 GGGCAGGGATGGCGATCGCGGGCGAGATCAAGCTTCCACCGTTTCATATCGGGCCTACC 141
Qy 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
Db 142 GCCTTCCTCGGCTGGGTGTTGTCGACAAACACGCAACGCGCGAGTCCACGCGTG 201
Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
Db 202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGACGCTATCACCGCGGTC 261
Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100
Db 262 GACGGCGCTCCGATCAACTCGGCCACCGCATGGCGGACGCGCTTAACGGGCATCATCCC 321
```

```
Qy 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
Db 322 GGTGACGTCTCTCGGTGACCTGGCAACCAAGTCGGCGCGCACGCTACAGGACGCTG 381
Qy 121 ThrLeuAlaGluGlyProProAla 128
Db 382 ACATTGGCGGAGGACCCCGGCC 405

RESULT 5
US-09-643-597-351
; Sequence 351, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Hosken, Nancy
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643,597
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 351
; LENGTH: 1012
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-643-597-351

Alignment Scores:
Pred. No.: 2,49e-64 Length: 1012
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-684-215A-18 (1-128) x US-09-643-597-351 (1-1012)

Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20
Db 67 ACGGCGCGCTCCGATAACTTCCAGCTGTCCAGGGTGGCGAGGATTGCCATTCCCGATC 126
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
Db 127 GGGCAGGGATGGCGATCGCGGGCGAGATCAAGCTTCCACCGTTTCATATCGGGCCTACC 186
Qy 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
Db 187 GCCTTCCTCGGCTGGGTGTTGTCGACAAACACGCAACGCGCGAGTCCACGCGCTG 246
Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
Db 247 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGACGCTGATCACCGCGTC 306
Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100
Db 307 GACGGCGCTCCGATCAACTCGGCCACCGCATGGCGGACGCGCTTAACGGGCATCATCCC 366
Qy 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
Db 367 GGTGACGTCTCTCGGTGACCTGGCAACCAAGTCGGCGCGCACGCTACAGGACGCTG 426
Qy 121 ThrLeuAlaGluGlyProProAla 128
```

Db 427 ACATTGGCGAGGACCCCGGCC 450

## RESULT 6

US-09-606-421B-351

; Sequence 351, Application US/09606421B

; Patent No. 6531315

; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong

; APPLICANT: Fan, Liqun

; APPLICANT: Kalos, Michael D.

; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: Hosken, Nancy

; APPLICANT: Fanger, Gary R.

; APPLICANT: Li, Samuel X.

; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yasir A.W.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE REFERENCE: AND DIAGNOSIS OF LUNG CANCER

; CURRENT APPLICATION NUMBER: US/09/606,421B

; FILE REFERENCE: 210121.455C9

; CURRENT FILING DATE: 2000-06-28

; NUMBER OF SEQ ID NOS: 358

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 351

; LENGTH: 1012

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-606-421B-351

Alignment Scores:

Pred. No.: 2,49e-64 Length: 1012  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
Indels: 0  
Gaps: 4

US-09-684-215A-18 (1-128) x US-09-606-421B-351 (1-1012)

QY 1 ThrAlaAlaSerAspAsnPhelGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20

|||||

Db 67 ACGGCGCGTCCGATTAACCTCCAGCTGTCCAGGTGGCAGGGATTCGCCATTCGGATC 126

|||||

QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40

|||||

Db 127 GGGCAGCGCATCGGCGCCAGATCAAGCTTCCACCGTTTCATATCGGGCTACC 186

|||||

QY 41 AlaPheLeuGlyLeuGlyValValAlaAspAsnGlyAsnGlyAlaArgValGlnArgVal 60

|||||

Db 187 GCCTTCCTCGGCTGGGTGTTCGACAAACACGCGCAGCGGCGCAGGATCCCAACGCGTG 246

|||||

QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80

|||||

Db 247 GTCGGAGCGCTCCGCGCGCAAGTCTCGCATCTCCACCGCGGCGGATCCACCGGTC 306

|||||

QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100

|||||

Db 307 GACGGCGCTCCGATCAACTCGCGCCACCGGATGGCGGCGCGCTTAACGGCGCATCC 366

|||||

QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120

|||||

Db 367 GTGACGCTCATCTCGGTGACCTGGCAACCAAGTCCGGCGCGCAGCGGTACAGGACG 426

|||||

QY 121 ThrLeuAlaGluGlyProProAla 128

|||||

Db 427 ACATTGGCGAGGACCCCGGCC 450

|||||

## RESULT 7

US-09-620-412C-348

; Sequence 348, Application US/09620412C

; Patent No. 6448234

; GENERAL INFORMATION:

; APPLICANT: Steven P. Fling  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
; FILE REFERENCE: DIAGNOSIS OF CHLAMYDIAL INFECTION

; CURRENT APPLICATION NUMBER: US/09/620,412C

; CURRENT FILING DATE: 2000-07-20

; NUMBER OF SEQ ID NOS: 363

; SOFTWARE: FastSeq for Windows Version 3.0/4.0

; SEQ ID NO 348

; LENGTH: 1464

; TYPE: DNA

; ORGANISM: Chlamydia trachomatis

US-09-620-412C-348

Alignment Scores:

Pred. No.: 4,06e-64 Length: 1464  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
Indels: 0  
Gaps: 4

US-09-684-215A-18 (1-128) x US-09-620-412C-348 (1-1464)

QY 1 ThrAlaAlaSerAspAsnPhelGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20

|||||

Db 22 ACGGCGCGTCCGATTAACCTCCAGCTGTCCAGGTGGCAGGGATTCGCCATTCGGATC 81

|||||

QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40

|||||

Db 82 GGGCAGCGCATCGGCGCCAGATCAAGCTTCCACCGTTTCATATCGGGCTACC 141

|||||

QY 41 AlaPheLeuGlyLeuGlyValValAlaAspAsnGlyAsnGlyAlaArgValGlnArgVal 60

|||||

Db 142 GCCTTCCTCGGCTGGGTGTTCGACAAACACGCGCAGCGGATCCCAACGCGTG 201

|||||

QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80

|||||

Db 202 GTCGGAGCGCTCCGCGCGCAAGTCTCGGCATCTCCACCGCGGCGGATCCACCGGTC 261

|||||

QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100

|||||

Db 262 GACGGCGCTCCGATCAACTCGCGCCACCGGATGGCGGCGCGCTTAACGGCGCATCC 321

|||||

QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120

|||||

Db 322 GTGACGCTCATCTCGGTGACCTGGCAACCAAGTCCGGCGCGCAGCGGTACAGGACG 381

|||||

QY 121 ThrLeuAlaGluGlyProProAla 128

|||||

Db 382 ACATTGGCGAGGACCCCGGCC 405

|||||

## RESULT 8

US-09-598-419-348

; Sequence 348, Application US/09598419

; Patent No. 6565856

; GENERAL INFORMATION:

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Scholler, John

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND

; FILE REFERENCE: DIAGNOSIS OF CHLAMYDIAL INFECTION

; CURRENT APPLICATION NUMBER: US/09/598,419

; NUMBER OF SEQ ID NOS: 357

; SOFTWARE: FastSeq for Windows Version 3.0/4.0

; SEQ ID NO 348

; LENGTH: 1464

; TYPE: DNA

; ORGANISM: Chlamydia trachomatis

US-09-598-419-348

Alignment Scores:

Pred. No.: 4,06e-64 Length: 1464  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-684-215A-18 (1-128) x US-09-598-419-348 (1-1464)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
|||||  
DB 22 AGGGCCGCTCCGATTAACCTCCAGCTGTCCAGGGTGGCAGGGATTCGCCATTCGGATC 81  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
|||||  
DB 82 GGCAGCGCATGGCGATCGGGCCAGATCAAGCTTCCACCGCTTCATATCGGGCTACC 141  
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60  
|||||  
DB 142 GCCTTCCTCGGCTGGGTGTTCGACAAACAGGCAACGGCGACGATCCCAACCGGTG 201  
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
|||||  
DB 202 GTCCGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGGGCAGCTGATCACCAGCGGTG 261  
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAlaAspAlaLeuAsnGlyHisPro 100  
|||||  
DB 262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCC 321  
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
|||||  
DB 322 GGTGAGCTCATCTCGGTGACCTGGCAACCAAGTCGGGGCGGACGCTACAGGGAACGTG 381  
QY 121 ThrLeuAlaGluGlyProAla 128  
|||||  
DB 382 ACATTGGCGAGGGACCCCGGCC 405

## RESULT 9

US-09-620-412C-332  
; Sequence 332, Application US/09620412C  
; Patent No. 6448234

; GENERAL INFORMATION:  
; APPLICANT: Steven P. Fling  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
; FILE REFERENCE: DIAGNOSIS OF CHLAMYDIAL INFECTION  
; FILE REFERENCE: 210121.469C7  
; CURRENT APPLICATION NUMBER: US/09/620,412C  
; CURRENT FILING DATE: 2000-07-20  
; NUMBER OF SEQ ID NOS: 363

; SOFTWARE: FastSeq for Windows Version 3.0/4.0  
; SEQ ID NO 332

; LENGTH: 1557

; TYPE: DNA

; ORGANISM: Chlamydia trachomatis

US-09-620-412C-332

Alignment Scores:  
Pred. No.: 4,4e-64 Length: 1557  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-684-215A-18 (1-128) x US-09-620-412C-332 (1-1557)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
|||||  
DB 22 AGGGCCGCTCCGATTAACCTCCAGCTGTCCAGGGTGGCAGGGATTCGCCATTCGGATC 81  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
|||||  
DB 82 GGCAGCGCATGGCGATCGGGCCAGATCAAGCTTCCACCGCTTCATATCGGGCTACC 141

QY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60  
|||||  
DB 142 GCCTTCCTCGGCTGGGTGTTCGACAAACAGGCAACGGCGACGATCCCAACCGGTG 201  
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
|||||  
DB 202 GTCCGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGGGCAGCTGATCACCAGCGGTG 261  
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAlaAspAlaLeuAsnGlyHisPro 100  
|||||  
DB 262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCC 321  
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
|||||  
DB 322 GGTGAGCTCATCTCGGTGACCTGGCAACCAAGTCGGGGCGGACGCTACAGGGAACGTG 381  
QY 121 ThrLeuAlaGluGlyProAla 128  
|||||  
DB 382 ACATTGGCGAGGGACCCCGGCC 405

## RESULT 10

US-09-598-419-332

; Sequence 332, Application US/09598419

; Patent No. 6565856

; GENERAL INFORMATION:

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Scholler, John

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND

; FILE REFERENCE: DIAGNOSIS OF CHLAMYDIAL INFECTION

; FILE REFERENCE: 210121.469C6

; CURRENT APPLICATION NUMBER: US/09/598,419

; CURRENT FILING DATE: 2000-06-20

; NUMBER OF SEQ ID NOS: 357

; SOFTWARE: FastSeq for Windows Version 3.0/4.0

; SEQ ID NO 332

; LENGTH: 1557

; TYPE: DNA

; ORGANISM: Chlamydia trachomatis

US-09-598-419-332

Alignment Scores:  
Pred. No.: 4,4e-64 Length: 1557  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-684-215A-18 (1-128) x US-09-598-419-332 (1-1557)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
|||||  
DB 22 AGGGCCGCTCCGATTAACCTCCAGCTGTCCAGGGTGGCAGGGATTCGCCATTCGGATC 81  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
|||||  
DB 82 GGCAGCGCATGGCGATCGGGCCAGATCAAGCTTCCACCGCTTCATATCGGGCTACC 141  
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60  
|||||  
DB 142 GCCTTCCTCGGCTGGGTGTTCGACAAACAGGCAACGGCGACGATCCCAACCGGTG 201  
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
|||||  
DB 202 GTCCGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGGGCAGCTGATCACCAGCGGTG 261  
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAlaAspAlaLeuAsnGlyHisPro 100  
|||||  
DB 262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCC 321  
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
|||||

Db 322 GGTGACGTATCTCGGTGACCTGGCAACCAAGTCGGCGGCGACCGTACAGGGAACGTG 381  
Qy 121 ThrLeuAlaGluGlyProAla 128  
|||||  
Db 382 ACATTGGCGGAGGACCCCGGCC 405

## RESULT 11

US-09-556-877-188

; Sequence 188, Application US/09556877

; Patent No. 6432916

; GENERAL INFORMATION:

; APPLICANT: Probst, Peter

; APPLICANT: Bhatia, Ajay

; APPLICANT: Skeiky, Yasir

; APPLICANT: Fling, Steve

; APPLICANT: Maisonneuve, Jeff

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND

; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION

; FILE REFERENCE: 210121.469C5

; CURRENT APPLICATION NUMBER: US/09/556,877

; NUMBER OF SEQ ID NOS: 305

; SOFTWARE: FastSeq for Windows Version 3.0/4.0

; SEQ ID NO 188

; LENGTH: 1578

; TYPE: DNA

; ORGANISM: Chlamydia

US-09-556-877-188

Alignment Scores:  
Pred. No.: 4,48e-64 Length: 1578  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-684-215A-18 (1-128) x US-09-556-877-188 (1-1578)

Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20  
|||||  
Db 22 ACGGCGCGCTCCGATAACTTCCAGCTGTCCAGGCTGGCGAGGATTCCGCATTCGGATC 81  
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLeuProThrValHisIleGlyProThr 40  
|||||  
Db 82 GGGCAGGCGATGGCGATCGCGGCGAGATCAAGCTTCCACCGTTTCATATCGGGCTACC 141  
Qy 41 AlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArgValGlnArgVal 60  
|||||  
Db 142 GCCTTCCTCGGCTGGGTGTGTCGACACACGACGCGGCGAGTCCACCGCTTCAATCGGGCTG 201  
Qy 61 ValGlySerAlaProAlaSerLeuGlyValValAspAsnGlyAsnGlyAlaArgValGlnArgVal 80  
|||||  
Db 142 GCCTTCCTCGGCTGGGTGTGTCGACACACGACGCGGCGAGTCCACCGCTTCAATCGGGCTG 201  
Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100  
|||||  
Db 262 GACGCGCTCCGATCAACTCGGCCACCGCGATGGCGGCGAGTCCACCGCTTCAATCGGGCTG 321  
Qy 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
|||||  
Db 322 GGTGACGTATCTCGGTGACCTGGCAACCAAGTCGGCGGCGACCGTACAGGGAACGTG 381  
Qy 121 ThrLeuAlaGluGlyProAla 128  
|||||  
Db 382 ACATTGGCGGAGGACCCCGGCC 405

## RESULT 12

US-09-620-412C-188

; Sequence 188, Application US/09620412C

; Patent No. 6448234

; GENERAL INFORMATION:

; APPLICANT: Steven P. Fling

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND

; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION

; FILE REFERENCE: 210121.469C7

; CURRENT APPLICATION NUMBER: US/09/620,412C

; CURRENT FILING DATE: 2000-07-20

; NUMBER OF SEQ ID NOS: 363

; SOFTWARE: FastSeq for Windows Version 3.0/4.0

; SEQ ID NO 188

; LENGTH: 1578

; TYPE: DNA

; ORGANISM: Chlamydia

US-09-620-412C-188

Alignment Scores:  
Pred. No.: 4,48e-64 Length: 1578  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-684-215A-18 (1-128) x US-09-620-412C-188 (1-1578)

Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20  
|||||  
Db 22 ACGGCGCGCTCCGATAACTTCCAGCTGTCCAGGCTGGCGAGGATTCCGCATTCGGATC 81  
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLeuProThrValHisIleGlyProThr 40  
|||||  
Db 82 GGGCAGGCGATGGCGATCGCGGCGAGATCAAGCTTCCACCGTTTCATATCGGGCTACC 141  
Qy 41 AlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArgValGlnArgVal 60  
|||||  
Db 142 GCCTTCCTCGGCTGGGTGTGTCGACACACGACGCGGCGAGTCCACCGCTTCAATCGGGCTG 201  
Qy 61 ValGlySerAlaProAlaSerLeuGlyValValAspAsnGlyAsnGlyAlaArgValGlnArgVal 80  
|||||  
Db 202 GTCGGGAGCGCTCCCGGCGCAAGTCTCGGCATCTCCACCGGCGAGTATCACCGCGGTG 261  
Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100  
|||||  
Db 262 GACGCGCTCCGATCAACTCGGCCACCGCGATGGCGGCGAGTCCACCGCTTCAATCGGGCTG 321  
Qy 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
|||||  
Db 322 GGTGACGTATCTCGGTGACCTGGCAACCAAGTCGGCGGCGACCGTACAGGGAACGTG 381  
Qy 121 ThrLeuAlaGluGlyProAla 128  
|||||  
Db 382 ACATTGGCGGAGGACCCCGGCC 405

## RESULT 13

US-09-598-419-188

; Sequence 188, Application US/09598419

; Patent No. 6565856

; GENERAL INFORMATION:

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Scholler, John

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND

; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION

; FILE REFERENCE: 210121.469C6

; CURRENT APPLICATION NUMBER: US/09/598,419

; CURRENT FILING DATE: 2000-06-20

; NUMBER OF SEQ ID NOS: 357

; SOFTWARE: FastSeq for Windows Version 3.0/4.0

; SEQ ID NO 188

; LENGTH: 1578

; TYPE: DNA

; ORGANISM: Chlamydia

US-09-598-419-188

Alignment Scores:

Pred. No.: 4,48e-64 Length: 1578



```
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-684-215A-18 (1-128) x US-09-598-419-188 (1-1578)

Qy 1 ThrAlaAlaSerAspAsnGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20
Db 22 ACGCGCGCTCCGATAACTTCCAGCTGTCCAGGGTGGCAGGATTCGCCATTCGATC 81
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
Db 82 GGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCACCGTTTCATATCGGGCCCTACC 141
Qy 41 AlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArgValGlnArgVal 60
Db 142 GCCTTCTCGGCTTGGGTGTGTCGACAAACAAACGCGACGCGGATTCACAGCGGTG 201
Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
Db 202 GTCCGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGCGACGTATCACCGCGGTC 261
Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisH1sPro 100
Db 262 GACGGCGCTCCGATCAACTCGGCCACCGCATGGGACGCGCTTAACGGGCATCATCCC 321
Qy 101 GlyAspValIleSerValThrTrrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
Db 322 GGTACGCTCATCTCGGTGACCTGGCAACCAACGACGCGGCGGCGGTACAGGGAACGTG 381
Qy 121 ThrLeuAlaGluGlyProAla 128
Db 382 ACATTGGCGAGGACCCCGGCC 405

RESULT 14
US-09-620-412C-352
; Sequence 352, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
; APPLICANT: Steven P. Fling
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C7
; CURRENT APPLICATION NUMBER: US/09/620.412C
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 352
; LENGTH: 1752
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis
US-09-620-412C-352

Alignment Scores:
Pred. No.: 5,14e-64 Length: 1752
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-684-215A-18 (1-128) x US-09-620-412C-352 (1-1752)

Qy 1 ThrAlaAlaSerAspAsnGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20
Db 22 ACGCGCGCTCCGATAACTTCCAGCTGTCCAGGGTGGCAGGATTCGCCATTCGATC 81
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
Db 82 GGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCACCGTTTCATATCGGGCCCTACC 141
Qy 41 AlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArgValGlnArgVal 60
Db 142 GCCTTCTCGGCTTGGGTGTGTCGACAAACAAACGCGACGCGGATTCACAGCGGTG 201
Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
Db 202 GTCCGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGCGACGTATCACCGCGGTC 261
Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisH1sPro 100
Db 262 GACGGCGCTCCGATCAACTCGGCCACCGCATGGGACGCGCTTAACGGGCATCATCCC 321
Qy 101 GlyAspValIleSerValThrTrrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
Db 322 GGTACGCTCATCTCGGTGACCTGGCAACCAACGACGCGGCGGCGGTACAGGGAACGTG 381
Qy 121 ThrLeuAlaGluGlyProAla 128
Db 382 ACATTGGCGAGGACCCCGGCC 405

RESULT 15
US-09-598-419-352
; Sequence 352, Application US/09598419
; Patent No. 6565856
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C6
; CURRENT APPLICATION NUMBER: US/09/598.419
; CURRENT FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 357
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 352
; LENGTH: 1752
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis
US-09-598-419-352

Alignment Scores:
Pred. No.: 5,14e-64 Length: 1752
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-684-215A-18 (1-128) x US-09-598-419-352 (1-1752)

Qy 1 ThrAlaAlaSerAspAsnGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20
Db 22 ACGCGCGCTCCGATAACTTCCAGCTGTCCAGGGTGGCAGGATTCGCCATTCGATC 81
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
Db 82 GGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCACCGTTTCATATCGGGCCCTACC 141
Qy 41 AlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArgValGlnArgVal 60
Db 142 GCCTTCTCGGCTTGGGTGTGTCGACAAACAAACGCGACGCGGATTCACAGCGGTG 201
Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
Db 202 GTCCGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGCGACGTATCACCGCGGTC 261
Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisH1sPro 100
Db 262 GACGGCGCTCCGATCAACTCGGCCACCGCATGGGACGCGCTTAACGGGCATCATCCC 321
Qy 101 GlyAspValIleSerValThrTrrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
Db 322 GGTACGCTCATCTCGGTGACCTGGCAACCAACGACGCGGCGGCGGTACAGGGAACGTG 381
```

```
Qy 41 AlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArgValGlnArgVal 60
Db 142 GCCTTCTCGGCTTGGGTGTGTCGACAAACAAACGCGACGCGGATTCACAGCGGTG 201
Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
Db 202 GTCCGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGCGACGTATCACCGCGGTC 261
Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisH1sPro 100
Db 262 GACGGCGCTCCGATCAACTCGGCCACCGCATGGGACGCGCTTAACGGGCATCATCCC 321
Qy 101 GlyAspValIleSerValThrTrrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
Db 322 GGTACGCTCATCTCGGTGACCTGGCAACCAACGACGCGGCGGCGGTACAGGGAACGTG 381
Qy 121 ThrLeuAlaGluGlyProAla 128
Db 382 ACATTGGCGAGGACCCCGGCC 405

RESULT 15
US-09-598-419-352
; Sequence 352, Application US/09598419
; Patent No. 6565856
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C6
; CURRENT APPLICATION NUMBER: US/09/598.419
; CURRENT FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 357
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 352
; LENGTH: 1752
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis
US-09-598-419-352

Alignment Scores:
Pred. No.: 5,14e-64 Length: 1752
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-684-215A-18 (1-128) x US-09-598-419-352 (1-1752)

Qy 1 ThrAlaAlaSerAspAsnGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20
Db 22 ACGCGCGCTCCGATAACTTCCAGCTGTCCAGGGTGGCAGGATTCGCCATTCGATC 81
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
Db 82 GGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCACCGTTTCATATCGGGCCCTACC 141
Qy 41 AlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArgValGlnArgVal 60
Db 142 GCCTTCTCGGCTTGGGTGTGTCGACAAACAAACGCGACGCGGATTCACAGCGGTG 201
Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
Db 202 GTCCGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGCGACGTATCACCGCGGTC 261
Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisH1sPro 100
Db 262 GACGGCGCTCCGATCAACTCGGCCACCGCATGGGACGCGCTTAACGGGCATCATCCC 321
Qy 101 GlyAspValIleSerValThrTrrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
Db 322 GGTACGCTCATCTCGGTGACCTGGCAACCAACGACGCGGCGGCGGTACAGGGAACGTG 381
```

Qy 121 ThrLeuAlaGluGlyProProAla 128  
|||||  
Db 382 ACATTGGCCGAGGGACCCCGGCC 405

Search completed: September 5, 2003, 12:12:10  
Job time : 75.4138 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 5, 2003, 07:52:39 ; Search time 2285.46 Seconds  
(without alignments)  
2291.192 Million cell updates/sec

Title: US-09-684-215A-18  
Perfect score: 653  
Sequence: 1 TAASDNFQSLGGGQFAIPI.....QTKSGGTRTGNVTLAEGPPA 128

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlp  
-O=/cn2\_1/USPTO\_spool\_p/US09684215/runat\_05092003\_072202\_559/app\_query.fasta\_1.853  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09684215.acgn\_1.1.3962 -runat\_05092003\_072202\_559 -NCPU=6 -ICPU=3  
-NO\_MMAL -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl :  
1: gb\_ba : \*  
2: gb\_htg : \*  
3: gb\_in : \*  
4: gb\_om : \*  
5: gb\_ov : \*  
6: gb\_pat : \*  
7: gb\_ph : \*  
8: gb\_pl : \*  
9: gb\_pr : \*  
10: gb\_ro : \*  
11: gb\_sts : \*  
12: gb\_sy : \*  
13: gb\_un : \*  
14: gb\_vi : \*  
15: em\_ba : \*  
16: em\_fun : \*  
17: em\_hum : \*  
18: em\_in : \*  
19: em\_mu : \*  
20: em\_om : \*  
21: em\_or : \*  
22: em\_ov : \*  
23: em\_pat : \*  
24: em\_ph : \*  
25: em\_pl : \*  
26: em\_ro : \*  
27: em\_sts : \*  
28: em\_un : \*

29: em\_vi : \*  
30: em\_htg\_hum : \*  
31: em\_htg\_inv : \*  
32: em\_htg\_other : \*  
33: em\_htg\_mus : \*  
34: em\_htg\_pln : \*  
35: em\_htg\_rod : \*  
36: em\_htg\_mam : \*  
37: em\_htg\_vrt : \*  
38: em\_sy : \*  
39: em\_htgo\_hum : \*  
40: em\_htgo\_mus : \*  
41: em\_htgo\_other : \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	653	100.0	675	6	AR261272 Sequence
2	653	100.0	675	6	AX201049 Sequence
3	653	100.0	675	6	AX267848 Sequence
4	653	100.0	822	6	AR277645 Sequence
5	653	100.0	822	6	AX369152 Sequence
6	653	100.0	894	6	AX351489 Sequence
7	653	100.0	900	6	AR220690 Sequence
8	653	100.0	900	6	AX365960 Sequence
9	653	100.0	915	6	AR261281 Sequence
10	653	100.0	915	6	AX201061 Sequence
11	653	100.0	915	6	AX267860 Sequence
12	653	100.0	945	6	AR277644 Sequence
13	653	100.0	945	6	AX369151 Sequence
14	653	100.0	1012	6	AR220689 Sequence
15	653	100.0	1012	6	AX365958 Sequence
16	653	100.0	1155	6	AX369165 Sequence
17	653	100.0	1203	6	AX201078 Sequence
18	653	100.0	1203	6	AX267877 Sequence
19	653	100.0	1464	6	AR229410 Sequence
20	653	100.0	1464	6	AX156105 Sequence
21	653	100.0	1464	6	AX361955 Sequence
22	653	100.0	1557	6	AR229398 Sequence
23	653	100.0	1557	6	AX156089 Sequence
24	653	100.0	1557	6	AX361939 Sequence
25	653	100.0	1578	6	AR229328 Sequence
26	653	100.0	1578	6	AX155945 Sequence
27	653	100.0	1578	6	AX361795 Sequence
28	653	100.0	1590	6	AX316986 Sequence
29	653	100.0	1752	6	AR229413 Sequence
30	653	100.0	1752	6	AX156109 Sequence
31	653	100.0	1752	6	AX361959 Sequence
32	653	100.0	1758	6	AR229401 Sequence
33	653	100.0	1758	6	AX156093 Sequence
34	653	100.0	1758	6	AX361943 Sequence
35	653	100.0	1860	6	AR229380 Sequence
36	653	100.0	1860	6	AX156065 Sequence
37	653	100.0	1860	6	AX361915 Sequence
38	653	100.0	1896	6	AR229392 Sequence
39	653	100.0	1896	6	AX156081 Sequence
40	653	100.0	1896	6	AX361931 Sequence
41	653	100.0	1941	6	AR229386 Sequence
42	653	100.0	1941	6	AX156073 Sequence
43	653	100.0	1941	6	AX361923 Sequence
44	653	100.0	1965	6	AR229404 Sequence
45	653	100.0	1965	6	AX156097 Sequence

ALIGNMENTS

RESULT 1

AR261272 LOCUS 675 bp DNA linear PAT 29-JAN-2003  
SEQUENCE 822 from patent US 6321716.  
ACCESSION AR261272  
VERSION AR261272.1 GI:28072035  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 675)  
AUTHORS Mashiki,Z. and Harada,J.  
TITLE Negative pressure control apparatus for engine mounted in vehicle  
JOURNAL Patent: US 6321716-A 822 27-NOV-2001;  
FEATURES Location/Qualifiers  
source 1..675  
BASE COUNT 162 a 197 c 190 g 126 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 2.28e-44 Length: 675  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-09-684-215A-18 (1-128) x AR261272 (1-675)  
QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
DB 22 ACAGCGCGCTCGATAACTTCCAGCTGTCCAGGGTGGCAGGGATTCGCCATTCCGATC 81  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
DB 82 GGGCAGGCATGGCGATCGCGGCCAGATCAAGCTTCCACCGCTTCATATCGGGCTACC 141  
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60  
DB 142 GCCTTCTCGCTTGGGTGTTCTGCACAAACACGACGCGGCGAGTTCACCGCGTG 201  
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
DB 202 GTCCGGAGCGCTCCCGCGCAAGTCTCGGCATCTCCACGGGCGAGTGATCACCGCGTC 261  
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100  
DB 262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCC 321  
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
DB 322 GGTGACGTATCTCGGTGACCTGGCAACCAAGTCGGCGGCACGCGTACAGGGAACGTG 381  
QY 121 ThrLeuAlaGluGlyProProAla 128  
DB 382 ACATTGGCGGAGGACCCCGGCC 405  
RESULT 2  
LOCUS AX201049 675 bp DNA linear PAT 29-AUG-2001  
DEFINITION Sequence 679 from Patent WO0151633.  
ACCESSION AX201049  
VERSION AX201049.1 GI:15390857  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1  
AUTHORS Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,  
Reed,S.G., Kalos,M.D., Fanger,G.R., Day,C.H., Retter,M.W.,  
Stolk,J.A., Skeiky,Y.A., Wang,A. and Meagher,M.J.  
TITLE Compositions and methods for the therapy and diagnosis of prostate  
cancer  
Patent: WO 0173032-A 822 04-OCT-2001;  
CORIXA CORPORATION (US)  
Location/Qualifiers  
source 1..675  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

cancer  
Patent: WO 0151633-A 679 19-JUL-2001;  
CORIXA CORPORATION (US)  
Location/Qualifiers  
source 1..675  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

BASE COUNT 162 a 197 c 190 g 126 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 2.28e-44 Length: 675  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-09-684-215A-18 (1-128) x AX201049 (1-675)  
QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
DB 22 ACAGCGCGCTCGATAACTTCCAGCTGTCCAGGGTGGCAGGGATTCGCCATTCCGATC 81  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
DB 82 GGGCAGGCATGGCGATCGCGGCCAGATCAAGCTTCCACCGCTTCATATCGGGCTACC 141  
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60  
DB 142 GCCTTCTCGCTTGGGTGTTCTGCACAAACACGACGCGGCGAGTTCACCGCGTG 201  
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
DB 202 GTCCGGAGCGCTCCCGCGCAAGTCTCGGCATCTCCACGGGCGAGTGATCACCGCGTC 261  
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100  
DB 262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCC 321  
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
DB 322 GGTGACGTATCTCGGTGACCTGGCAACCAAGTCGGCGGCACGCGTACAGGGAACGTG 381  
QY 121 ThrLeuAlaGluGlyProProAla 128  
DB 382 ACATTGGCGGAGGACCCCGGCC 405  
RESULT 3  
LOCUS AX267848 675 bp DNA linear PAT 26-OCT-2001  
DEFINITION Sequence 822 from Patent WO0173032.  
ACCESSION AX267848  
VERSION AX267848.1 GI:16516494  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1  
AUTHORS Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,  
Kalos,M.D., Fanger,G.R., Retter,M.W., Stolk,J.A., Day,C.H.,  
Vedvick,T.S., Carter,D., Li,S.X., Wang,A., Skeiky,Y.A., Hepler,W.T.  
and Henderson,R.A.  
TITLE Compositions and methods for the therapy and diagnosis of prostate  
cancer  
Patent: WO 0173032-A 822 04-OCT-2001;  
CORIXA CORPORATION (US)  
Location/Qualifiers  
source 1..675  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"

BASE COUNT		162 a		197 c		190 g		126 t		/db_xref="taxon:9606"	
ORIGIN											
Alignment Scores:											
Pred. No.:		2.28e-44		Length:		675					
Score:		653.00		Matches:		128					
Percent Similarity:		100.00%		Conservative:		0					
Best Local Similarity:		100.00%		Mismatches:		0					
Query Match:		100.00%		Indels:		0					
DB:		6		Gaps:		0					
US-09-684-215A-18 (1-128) x AX267848 (1-675)											
Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyGlnGlyPheAlaIleProIle 20											
Db 22 ACGGCGCGCTCCGATAAATTCACAGTGTCCAGGCTGGCGAGGATTCGCCATTCGCATC 81											
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40											
Db 82 GGGCAGGCGATGGCGATCGCGGCGAGATCAAGCTTCCACCGTTCATATCGGGCGCTACC 141											
Qy 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60											
Db 142 GCCTTCCTCGGCTGGGTGTTCGACAAACACGCGGACGCGGACGAGTCCAAACGGGTG 201											
Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80											
Db 202 GTCGGAGCGCTCCGCGGCAAGTCTCGGCATCTCCACGCGGACGCTGATCACCGCGTACC 261											
Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100											
Db 262 GACGGCGCTCCGATCAACTCGGCCACCGCATGCGGACGCGTAAACGGCGATCATCCC 321											
Qy 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120											
Db 322 GGTGACGTATCTCGGTGACCTGGCAACCAAGTCGGCGGCGACGCGTACAGGAAACGTG 381											
Qy 121 ThrLeuAlaGluGlyProProAla 128											
Db 382 ACATTGGCGAGGACCCCGGCC 405											
RESULT 5											
LOCUS		AX369152		822 bp		DNA		linear		PAT 16-FEB-2002	
DEFINITION		Sequence 1862 from Patent WO0204514.									
ACCESSION		AX369152									
VERSION		AX369152.1 GI:18857170									
KEYWORDS		Homo sapiens (human)									
SOURCE		Homo sapiens									
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
REFERENCE											
AUTHORS		Wang, T., Watanabe, Y., Henderson, R.A., Johnson, J.C., Retter, M.W., Marnerakis, M., Carter, D., Fanger, G.R., Vedvick, T.S., Bangur, C.S., McNabb, A., Fanger, N., Switzer, A., McNeill, P.D. and Clapper, J.D.									
TITLE		Compositions and methods for the therapy and diagnosis of lung cancer									
JOURNAL		Patent: WO 0204514-A 1862 17-JAN-2002;									
FEATURES		CORIXA CORPORATION (US)									
source		1..822									
BASE COUNT		145 a		281 c		270 g		125 t			
ORIGIN											
Alignment Scores:											
Pred. No.:		2.79e-44		Length:		822					
Score:		653.00		Matches:		128					
Percent Similarity:		100.00%		Conservative:		0					
Best Local Similarity:		100.00%		Mismatches:		0					
Query Match:		100.00%		Indels:		0					
DB:		6		Gaps:		0					
US-09-684-215A-18 (1-128) x AX369152 (1-822)											
Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyGlnGlyPheAlaIleProIle 20											
Db 22 ACGGCGCGCTCCGATAAATTCACAGTGTCCAGGCTGGCGAGGATTCGCCATTCGCATC 81											
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40											
Db 82 GGGCAGGCGATGGCGATCGCGGCGAGATCAAGCTTCCACCGTTCATATCGGGCGCTACC 141											
Qy 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60											
Db 142 GCCTTCCTCGGCTGGGTGTTCGACAAACACGCGGACGCGGACGAGTCCAAACGGGTG 201											
Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80											
Db 202 GTCGGAGCGCTCCGCGGCAAGTCTCGGCATCTCCACGCGGACGCTGATCACCGCGTACC 261											
Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100											
Db 262 GACGGCGCTCCGATCAACTCGGCCACCGCATGCGGACGCGTAAACGGCGATCATCCC 321											
Qy 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120											
Db 322 GGTGACGTATCTCGGTGACCTGGCAACCAAGTCGGCGGCGACGCGTACAGGAAACGTG 381											
Qy 121 ThrLeuAlaGluGlyProProAla 128											
Db 382 ACATTGGCGAGGACCCCGGCC 405											
RESULT 4											
LOCUS		AR277645		822 bp		DNA		linear		PAT 10-APR-2003	
DEFINITION		Sequence 1862 from patent US 6509448.									
ACCESSION		AR277645									
VERSION		AR277645.1 GI:29711294									
KEYWORDS		Unknown.									
SOURCE		Unknown.									
ORGANISM		Unclassified.									
REFERENCE		1 (bases 1 to 822)									
AUTHORS		Wang, T., Bangur, C.S., Lodes, M.J., Fanger, G.R., Vedvick, T.S., Carter, D., Retter, M.W., Mannion, J., Fan, L. and Wang, A.									
TITLE		Compositions and methods for the therapy and diagnosis of lung cancer									
JOURNAL		Patent: US 6509448-A 1862 21-JAN-2003;									
FEATURES		Location/Qualifiers									
source		1..822									
BASE COUNT		146 a		281 c		270 g		125 t			
ORIGIN											
Alignment Scores:											
Pred. No.:		2.79e-44		Length:		822					
Score:		653.00		Matches:		128					
Percent Similarity:		100.00%		Conservative:		0					
Best Local Similarity:		100.00%		Mismatches:		0					
Query Match:		100.00%		Indels:		0					
DB:		6		Gaps:		0					
US-09-684-215A-18 (1-128) x AR277645 (1-822)											

```
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyValAlaArgValGlnArgVal 60
|||||
Db 142 GCCTTCTCGGCTGGGTGCTGCGACACACAGCGGCAACGGCGCATCCATCC 201
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
|||||
Db 202 GTCGGAGGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGGGACGTATCACC 261
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100
|||||
Db 262 GACGGCGCTCCGATCAACTCGGCGCACCGGATGGCGGCGCTTAACGGGCATCATCC 321
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
|||||
Db 322 GGTGACGTATCTCGGTGACCTGGCAACCAAGTCGGGGCGGACGGGTACAGGAA 381
QY 121 ThrLeuAlaGluGlyProProAla 128
|||||
Db 382 ACATTGGCGAGGGACCCCGGCC 405

RESULT 6
AX351489
LOCUS AX351489 894 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 236 from Patent WO0196390.
ACCESSION AX351489
VERSION AX351489.1 GI:18616835
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Jiang,Y., Hepler,W.T., Clapper,J.D., Wang,A. and Secrist,H.
TITLE Compositions and methods for the therapy and diagnosis of colon
JOURNAL Patent: WO 0196390-A 236 20-DEC-2001;
CORIXA CORPORATION (US)
FEATURES
source 1..894
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 207 a 269 c 237 g 181 t
ORIGIN
Alignment Scores:
Pred. No.: 3,05e-44 Length: 894
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-684-215A-18 (1-128) x AX351489 (1-894)
QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
|||||
Db 22 ACGGCGCGCTCCGATAAATTCACGCTGTCCAGGTGGCGGAGGATTCGCCATTCGATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
|||||
Db 82 GGGCAGCGCATCGGATCAACTTCACGCTGTCCAGGTGGCGGAGGATTCGCCATTCGATC 141
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyValAlaArgValGlnArgVal 60
|||||
Db 142 GCCTTCTCGGCTGGGTGCTGCGACACACAGCGGACGGCGGATTCATATCGGGCTAC 141
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
|||||
Db 202 GTCGGAGGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGGGACGTATCACC 261
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100
|||||
Db 262 GACGGCGCTCCGATCAACTCGGCGCACCGGATGGCGGCGCTTAACGGGCATCATCC 321
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
|||||
Db 322 GGTGACGTATCTCGGTGACCTGGCAACCAAGTCGGGGCGGACGGGTACAGGAA 381
QY 121 ThrLeuAlaGluGlyProProAla 128
|||||
Db 382 ACATTGGCGAGGGACCCCGGCC 405

RESULT 7
AX351489
LOCUS AX351489 900 bp DNA linear PAT 26-SEP-2002
DEFINITION Sequence 353 from patent US 6436072.
ACCESSION AR220690
VERSION AR220690.1 GI:23327471
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 900)
AUTHORS Wang,T., Fan,L., Kalos,M.D., Bangur,C.S., Hosken,N.A., Fanger,G.R.,
Li,S.X., Wang,A., Skeiky,Y.A.W., Henderson,R.A. and McNeill,P.D.
TITLE Compositions and methods for the therapy and diagnosis of lung
JOURNAL Patent: US 6436072-A 353 30-JUL-2002;
FEATURES
source 1..900
/organism="unknown"
BASE COUNT 219 a 239 c 246 g 196 t
ORIGIN
Alignment Scores:
Pred. No.: 3,07e-44 Length: 900
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-684-215A-18 (1-128) x AR220690 (1-900)
QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
|||||
Db 22 ACGGCGCGCTCCGATAAATTCACGCTGTCCAGGTGGCGGAGGATTCGCCATTCGATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
|||||
Db 82 GGGCAGCGCATCGGATCAACTTCACGCTGTCCAGGTGGCGGAGGATTCGCCATTCGATC 141
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyValAlaArgValGlnArgVal 60
|||||
Db 142 GCCTTCTCGGCTGGGTGCTGCGACACACAGCGGACGGCGGATTCATATCGGGCTAC 201
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
|||||
Db 202 GTCGGAGGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGGGACGTATCACC 261
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100
|||||
Db 262 GACGGCGCTCCGATCAACTCGGCGCACCGGATGGCGGCGCTTAACGGGCATCATCC 321
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
|||||
Db 322 GGTGACGTATCTCGGTGACCTGGCAACCAAGTCGGGGCGGACGGGTACAGGAA 381
QY 121 ThrLeuAlaGluGlyProProAla 128
|||||
Db 382 ACATTGGCGAGGGACCCCGGCC 405

RESULT 8
AX365960
```

```
Db 262 GACGGCGCTCCGATCAACTCGGCGCACCGGATGGCGGCGCTTAACGGGCATCATCC 321
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
|||||
Db 322 GGTGACGTATCTCGGTGACCTGGCAACCAAGTCGGGGCGGACGGGTACAGGAA 381
QY 121 ThrLeuAlaGluGlyProProAla 128
|||||
Db 382 ACATTGGCGAGGGACCCCGGCC 405

RESULT 7
AX220690
LOCUS AR220690 900 bp DNA linear PAT 26-SEP-2002
DEFINITION Sequence 353 from patent US 6436072.
ACCESSION AR220690
VERSION AR220690.1 GI:23327471
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 900)
AUTHORS Wang,T., Fan,L., Kalos,M.D., Bangur,C.S., Hosken,N.A., Fanger,G.R.,
Li,S.X., Wang,A., Skeiky,Y.A.W., Henderson,R.A. and McNeill,P.D.
TITLE Compositions and methods for the therapy and diagnosis of lung
JOURNAL Patent: US 6436072-A 353 30-JUL-2002;
FEATURES
source 1..900
/organism="unknown"
BASE COUNT 219 a 239 c 246 g 196 t
ORIGIN
Alignment Scores:
Pred. No.: 3,07e-44 Length: 900
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-684-215A-18 (1-128) x AR220690 (1-900)
QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
|||||
Db 22 ACGGCGCGCTCCGATAAATTCACGCTGTCCAGGTGGCGGAGGATTCGCCATTCGATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
|||||
Db 82 GGGCAGCGCATCGGATCAACTTCACGCTGTCCAGGTGGCGGAGGATTCGCCATTCGATC 141
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyValAlaArgValGlnArgVal 60
|||||
Db 142 GCCTTCTCGGCTGGGTGCTGCGACACACAGCGGACGGCGGATTCATATCGGGCTAC 201
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
|||||
Db 202 GTCGGAGGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGGGACGTATCACC 261
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100
|||||
Db 262 GACGGCGCTCCGATCAACTCGGCGCACCGGATGGCGGCGCTTAACGGGCATCATCC 321
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
|||||
Db 322 GGTGACGTATCTCGGTGACCTGGCAACCAAGTCGGGGCGGACGGGTACAGGAA 381
QY 121 ThrLeuAlaGluGlyProProAla 128
|||||
Db 382 ACATTGGCGAGGGACCCCGGCC 405

RESULT 8
AX365960
```

LOCUS AX365960 900 bp DNA linear PAT 15-FEB-2002  
DEFINITION Sequence 353 from Patent WO200174.  
ACCESSION AX365960  
VERSION AX365960.1 GI:18697458  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE 1  
AUTHORS Wang, T., Wang, A., Skeiky, Y.A., Li, S.X., Kalos, M.D., Henderson, R.A., McNeill, P.D., Carter, N., Retter, M.W., Marnerakis, M., Fanger, G.R., Vedwick, T.S., Carter, D., Watanabe, Y. and Peckham, D.W.  
TITLE Compositions and methods for the therapy and diagnosis of lung cancer  
JOURNAL Patent: WO 0200174-A 353 03-JAN-2002;  
CORIXA CORPORATION (US)  
FEATURES  
source Location/Qualifiers  
1 .900  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 219 a 239 c 246 g 196 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 3,07e-44 Length: 900  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6  
US-09-684-215A-18 (1-128) x AX365960 (1-900)  
QY 1 ThrAlaAlaSerAspAsnSerPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
Db 22 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGGTGGCAGGGATTGCCATTCCGATC 81  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
Db 82 GGCAGCGCATGGCGATGCGGGCCAGATCAAGCTTCCACCGCTTCATATCGGGCTACC 141  
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60  
Db 142 GCCTTCTCGGCTTGGTGTGTCGACAAACAGCGCAGCGGATGCGGCGCATCATCC 201  
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
Db 142 GCCTTCTCGGCTTGGTGTGTCGACAAACAGCGCAGCGGATGCGGCGCATCATCC 201  
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
Db 202 GTCCGAGGCGCTCCGCGCAAGTCTCCGATCTCCACCGCGCACGTGATCACC GGTC 261  
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAlaAspAlaLeuAsnGlyHisPro 100  
Db 262 GACGGCGCTCCGATCAACTCCAGCCACCGGATGGCGAGCGGCTTAACGGGCATCATCC 321  
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
Db 322 GGTGACGTATCTCGTGACCTGGCAACCAAGTCGGGGCGCACGGGTACAGGGAACGTG 381  
QY 121 ThrLeuAlaGluGlyProProAla 128  
Db 382 ACATTGGCGAGGACCCCGGCC 405  
RESULT 9  
AR261281  
LOCUS AR261281 915 bp DNA linear PAT 29-JAN-2003  
DEFINITION Sequence 834 from patent US 6321716.  
ACCESSION AR261281  
VERSION AR261281.1 GI:28072044  
KEYWORDS Unknown.  
SOURCE Unknown.  
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 915)  
AUTHORS Mashiki, Z. and Harada, J.  
TITLE Negative pressure control apparatus for engine mounted in vehicle  
JOURNAL Patent: US 6321716-A 834 27-NOV-2001;  
FEATURES  
source Location/Qualifiers  
1 .915  
/organism="unknown"  
BASE COUNT 167 a 280 c 284 g 184 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 3.12e-44 Length: 915  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6  
US-09-684-215A-18 (1-128) x AR261281 (1-915)  
QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
Db 22 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGGTGGCAGGGATTGCCATTCCGATC 81  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
Db 82 GGCAGCGCATGGCGATGCGGGCCAGATCAAGCTTCCACCGCTTCATATCGGGCTACC 141  
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60  
Db 142 GCCTTCTCGGCTTGGTGTGTCGACAAACAGCGCAGCGGATGCGGCGCATCATCC 201  
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
Db 202 GTCCGAGGCGCTCCGCGCAAGTCTCCGATCTCCACCGCGCACGTGATCACC GGTC 261  
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAlaAspAlaLeuAsnGlyHisPro 100  
Db 262 GACGGCGCTCCGATCAACTCCAGCCACCGGATGGCGAGCGGCTTAACGGGCATCATCC 321  
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
Db 322 GGTGACGTATCTCGTGACCTGGCAACCAAGTCGGGGCGCACGGGTACAGGGAACGTG 381  
QY 121 ThrLeuAlaGluGlyProProAla 128  
Db 382 ACATTGGCGAGGACCCCGGCC 405  
RESULT 10  
AX201061  
LOCUS AX201061 915 bp DNA linear PAT 29-AUG-2001  
DEFINITION Sequence 691 from Patent WO0151633.  
ACCESSION AX201061  
VERSION AX201061.1 GI:15390868  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE 1  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
TITLE Compositions and methods for the therapy and diagnosis of prostate cancer  
JOURNAL Patent: WO 0151633-A 691 19-JUL-2001;  
CORIXA CORPORATION (US)  
FEATURES  
source Location/Qualifiers  
1 .915  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 167 a 280 c 284 g 184 t

```
ORIGIN
Alignment Scores:
Pred. No.: 3.12e-44 Length: 915
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-684-215A-18 (1-128) x AX201061 (1-915)
QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyGlnGlyPheAlaIleProIle 20
DB 22 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGTTGGCGAGGATTCGCCATTCGGATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
DB 82 GGGCAGGCGATGGCGATCGGGCCAGATCAAGCTTCCACCGTTTCATATCGGGCTACC 141
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
DB 142 GCCTTCTCGGCTTGGGTGTTGCGACAACACGCGCACGCGGAGTCCACGCGCTG 201
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
DB 202 GTGCGGAGCGTCCGCGGCAAGTCTCGGCATCTCCACCGCGACGTATCACCGCGTC 261
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100
DB 262 GACGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGATCATCCC 321
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
DB 322 GGTGACGTCATCTCGTGACCTGGCAACCAAGTCGGCGCGCACGCGTACAGGGAACGTG 381
QY 121 ThrLeuAlaGluGlyProProAla 128
DB 382 ACATTGGCGAGGACCGCGGCC 405

RESULT 12
AX267860
LOCUS AX267860 915 bp DNA linear PAT 26-OCT-2001
DEFINITION Sequence 834 from Patent WO0173032.
ACCESSION AX267860
VERSION AX267860.1 GI:16516503
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
Kalos,M.D., Fanger,G.R., Retter,M.W., Stolk,J.A., Day,C.H.,
Vedvick,T.S., Carter,D., Li,S.X., Wang,A., Skeiky,Y.A., Hepler,W.T.
and Henderson,R.A.
TITLE Compositions and methods for the therapy and diagnosis of prostate
cancer
JOURNAL Patent: WO 0173032-A 834 04-OCT-2001;
CORIXA CORPORATION (US)
FEATURES
source
Location/Qualifiers
1..915
/organism="Homo sapiens"
/db_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 167 a 280 c 284 g 184 t
ORIGIN
Alignment Scores:
Pred. No.: 3.12e-44 Length: 915
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-684-215A-18 (1-128) x AX201061 (1-915)
QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyGlnGlyPheAlaIleProIle 20
DB 22 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGTTGGCGAGGATTCGCCATTCGGATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
DB 82 GGGCAGGCGATGGCGATCGGGCCAGATCAAGCTTCCACCGTTTCATATCGGGCTACC 141
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
DB 142 GCCTTCTCGGCTTGGGTGTTGCGACAACACGCGCACGCGGAGTCCACGCGCTG 201
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
DB 202 GTGCGGAGCGTCCGCGGCAAGTCTCGGCATCTCCACCGCGACGTATCACCGCGTC 261
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100
DB 262 GACGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGATCATCCC 321
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
DB 322 GGTGACGTCATCTCGTGACCTGGCAACCAAGTCGGCGCGCACGCGTACAGGGAACGTG 381
QY 121 ThrLeuAlaGluGlyProProAla 128
DB 382 ACATTGGCGAGGACCGCGGCC 405

RESULT 11
AX267860
LOCUS AX267860 915 bp DNA linear PAT 26-OCT-2001
DEFINITION Sequence 834 from Patent WO0173032.
ACCESSION AX267860
VERSION AX267860.1 GI:16516503
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
Kalos,M.D., Fanger,G.R., Retter,M.W., Stolk,J.A., Day,C.H.,
Vedvick,T.S., Carter,D., Li,S.X., Wang,A., Skeiky,Y.A., Hepler,W.T.
and Henderson,R.A.
TITLE Compositions and methods for the therapy and diagnosis of prostate
cancer
JOURNAL Patent: WO 0173032-A 834 04-OCT-2001;
CORIXA CORPORATION (US)
FEATURES
source
Location/Qualifiers
1..915
/organism="Homo sapiens"
/db_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 167 a 280 c 284 g 184 t
ORIGIN
Alignment Scores:
Pred. No.: 3.12e-44 Length: 915
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-684-215A-18 (1-128) x AX277644 (1-945)
QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyGlnGlyPheAlaIleProIle 20
DB 22 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGTTGGCGAGGATTCGCCATTCGGATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
DB 82 GGGCAGGCGATGGCGATCGGGCCAGATCAAGCTTCCACCGTTTCATATCGGGCTACC 141
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
DB 142 GCCTTCTCGGCTTGGGTGTTGCGACAACACGCGCACGCGGAGTCCACGCGCTG 201
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
DB 202 GTGCGGAGCGTCCGCGGCAAGTCTCGGCATCTCCACCGCGACGTATCACCGCGTC 261
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100
DB 262 GACGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGATCATCCC 321
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
DB 322 GGTGACGTCATCTCGTGACCTGGCAACCAAGTCGGCGCGCACGCGTACAGGGAACGTG 381
QY 121 ThrLeuAlaGluGlyProProAla 128
DB 382 ACATTGGCGAGGACCGCGGCC 405

RESULT 12
AX277644
LOCUS AX277644 945 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 1861 from patent US 6509448.
ACCESSION AX277644
VERSION AX277644.1 GI:29711293
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 945)
AUTHORS Wang,T., Bangur,C.S., Lodes,M.J., Fanger,G.R., Vedvick,T.S.,
Carter,D., Retter,M.W., Mannion,J., Fan,L. and Wang,A.
TITLE Compositions and methods for the therapy and diagnosis of lung
cancer
JOURNAL Patent: US 6509448-A 1861 21-JAN-2003;
FEATURES
Location/Qualifiers
1..945
/organism="unknown"
BASE COUNT 150 a 318 c 360 g 117 t
ORIGIN
Alignment Scores:
Pred. No.: 3.23e-44 Length: 945
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-684-215A-18 (1-128) x AR277644 (1-945)
QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyGlnGlyPheAlaIleProIle 20
DB 22 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGTTGGCGAGGATTCGCCATTCGGATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
DB 82 GGGCAGGCGATGGCGATCGGGCCAGATCAAGCTTCCACCGTTTCATATCGGGCTACC 141
```



```

Qy 41 AlapheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
Db 142 GCCTTCCTCGGCTGGGTGTGTCGACAAACAGCGGCAACGGCGAGTCCACGCGTG 201
Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
Db 202 GTCGGGAGCGCTCGGCGGCAAGTCTCGGATCTCCACCGCGGACGATGATCAGCGCGTG 261
Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
Db 262 GACGGCGCTCGGATCACTCGCCACCGCATGCGGAGCGGCTTAACGGGCATCATCCC 321
Qy 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
Db 322 GGTGACGTCTCTCGGTGACCTGGCAAAACCAAGTCGGGCGGCGACGCTACAGGGAACGTG 381
Qy 121 ThrLeuAlaGluGlyProProAla 128
Db 382 ACATTGGCCGAGGAGCCCCCGGCC 405
RESULT 13
LOCUS AX369151 945 bp DNA linear PAT 16-FEB-2002
DEFINITION Sequence 1861 from Patent WO0204514.
ACCESSION AX369151
VERSION AX369151.1 GI:18857169
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
REFERENCE
AUTHORS Wang,T., Watanabe,Y., Henderson,R.A., Johnson,J.C., Retter,M.W.,
Marnerakis,M., Carter,D., Fanger,G.R., Vedwick,T.S., Bangur,C.S.,
McNabb,A., Fanger,N., Switzer,A., McNeill,P.D. and Clapper,J.D.
TITLE Compositions and methods for the therapy and diagnosis of lung
cancer
JOURNAL Patent: WO 0204514-A 1861 17-JAN-2002;
CORIXA CORPORATION (US)
FEATURES
source 1. 945
/organism="Homo sapiens"
/db_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 150 a 318 c 360 g 117 t
ORIGIN
Alignment Scores:
Pred. No.: 3.23e-44 Length: 945
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-684-215A-18 (1-128) x AX369151 (1-945)
Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
Db 22 ACGCGCGCTCCGATACTTCCAGCTGTCCAGGTTGGCGGAGGATTCGCGCATTCGCGATC 81
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
Db 82 GGGCAGGCGATGGCGATCGCGGCGAGATCAAGCTTCCACCGTTCATATCGGCGCTACC 141
Qy 41 AlapheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
Db 142 GCCTTCCTCGGCTGGGTGTGTCGACAAACAGCGGCAACGGCGAGTCCACGCGTG 201
Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
Db 202 GTCGGGAGCGCTCGGCGGCAAGTCTCGGATCTCCACCGCGGACGATGATCAGCGCGTG 261

```

```

Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
Db 262 GACGGCGCTCGGATCACTCGCCACCGCATGCGGAGCGGCTTAACGGGCATCATCCC 321
Qy 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
Db 322 GGTGACGTCTCTCGGTGACCTGGCAAAACCAAGTCGGGCGGCGACGCTACAGGGAACGTG 381
Qy 121 ThrLeuAlaGluGlyProProAla 128
Db 382 ACATTGGCCGAGGAGCCCCCGGCC 405
RESULT 14
LOCUS AR220689 1012 bp DNA linear PAT 26-SEP-2002
DEFINITION Sequence 351 from patent US 6426072.
ACCESSION AR220689
VERSION AR220689.1 GI:23327470
KEYWORDS
SOURCE Unknown.
ORGANISM
REFERENCE
AUTHORS Wang,T., Fan,L., Kalos,M.D., Bangur,C.S., Hosken,N.A., Fanger,G.R.,
Li,S.X., Wang,A., Skeiky,Y.A.W., Henderson,R.A. and McNeill,P.D.
TITLE Compositions and methods for the therapy and diagnosis of lung
cancer
JOURNAL Patent: US 6426072-A 351 30-JUL-2002;
FEATURES
source Location/Qualifiers
1. 1012
/organism="unknown"
BASE COUNT 235 a 289 c 289 g 199 t
ORIGIN
Alignment Scores:
Pred. No.: 3.47e-44 Length: 1012
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-684-215A-18 (1-128) x AR220689 (1-1012)
Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
Db 67 ACGCGCGCTCCGATACTTCCAGCTGTCCAGGTTGGCGGAGGATTCGCGCATTCGCGATC 126
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
Db 127 GGGCAGGCGATGGCGATCGCGGCGAGATCAAGCTTCCACCGTTCATATCGGCGCTACC 186
Qy 41 AlapheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
Db 187 GCCTTCCTCGGCTGGGTGTGTCGACAAACAGCGGCAACGGCGAGTCCACGCGTG 246
Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
Db 247 GTCGGAGCGCTCGGCGGCAAGTCTCGGATCTCCACCGCGGACGATGATCAGCGCGTC 306
Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
Db 307 GACGGCGCTCGGATCACTCGCCACCGCGATGCGGAGCGGCTTAACGGGCATCATCCC 366
Qy 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
Db 367 GGTGACGTCTCTCGGTGACCTGGCAAAACCAAGTCGGGCGGCGACGCTACAGGGAACGTG 426
Qy 121 ThrLeuAlaGluGlyProProAla 128
Db 427 ACATTGGCCGAGGAGCCCCCGGCC 450
RESULT 15

```

AX365958  
 LOCUS AX365958 1012 bp DNA linear PAT 15-FEB-2002  
 DEFINITION Sequence 351 from Patent WO200174.  
 ACCESSION AX365958  
 VERSION AX365958.1 GI:18697457  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1  
 REFERENCE Wang, T., Wang, A., Skeiky, Y. A., Li, S. X., Kalos, M. D., Henderson, R. A.,  
 McNeill, P. D., Fanger, N., Retter, M. W., Marnerakis, M., Fanger, G. R.,  
 Vedwick, T. S., Carter, D., Watanabe, Y. and Peckham, D. W.  
 TITLE Compositions and methods for the therapy and diagnosis of lung  
 cancer  
 JOURNAL Patent: WO 0200174-A 351 03-JAN-2002;  
 CORIXA CORPORATION (US)  
 FEATURES  
 source  
 1..1012  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 BASE COUNT 235 a 289 c 289 g 199 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 3.47e-44 Length: 1012  
 Score: 653.00 Matches: 128  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0  
 US-09-684-215A-18 (1-128) x AX365958 (1-1012)  
 QY 1 ThrAlaAlaSerAspAsnPhenLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
 DB 67 ACGGGCGGCTCCGATAAATCCAGCTGTCCAGGTGGCGAGGATTCCGCATCCGATC 126  
 QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLeuProThrValHisIleGlyProThr 40  
 DB 127 GGGCAGGCGATGGCGATCGGGGCCAGATCAAGCTTCCACCGCTCATATCGGGCTACC 186  
 QY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60  
 DB 187 GCCTTCCCTCGGCTTGGTGTTCGACAACAACGCAACGGCGGACGAGTCCACGCGTG 246  
 QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
 DB 247 GTCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACGGCGAGTGATCACCGCGGTC 306  
 QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100  
 DB 307 GACGGCGCTCCGATCAACTCGGCCACCGGATGGCGGAGCGCTTAACGGGATCATCCC 366  
 QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyThrArgThrGlyAsnVal 120  
 DB 367 GGTGACGTATCTCGTGACCTGTGCAAAACCAAGTCGGCGGCGACGCTACAGGGAACGTG 426  
 QY 121 ThrLeuAlaGluGlyProAla 128  
 DB 427 ACATTGGCGGAGGAGACCCCGGCC 450

Search completed: September 5, 2003, 10:31:42  
 Job time : 2287.8 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 5, 2003, 08:27:24 ; Search time 2211.31 Seconds  
(without alignments)  
1406.846 Million cell updates/sec

Title: US-09-684-215A-18  
Perfect score: 653  
Sequence: 1 TAASDNFOLSQGGQGAIPA.....OTKSGGTRGNTVLAEGPPA 128

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame\_plus\_p2n.model -DEV=xl  
-Q/cgn2\_1/USPTO.spool\_p/US09684215/tunat\_05092003.072202.572/app\_query.fasta\_1.853  
-DB=EST -OPM=fastp -SUFFIX=est -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pcr -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFM=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09684215 -CGN\_1\_1\_4575\_tunat\_05092003.072202.572 -NCPU=3  
-NO\_MMAPP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:\*

- 1: em\_estba:\*
- 2: em\_esthum:\*
- 3: em\_estin:\*
- 4: em\_estnu:\*
- 5: em\_estov:\*
- 6: em\_estpl:\*
- 7: em\_estro:\*
- 8: em\_hic:\*
- 9: gb\_est1:\*
- 10: gb\_est2:\*
- 11: gb\_hic:\*
- 12: gb\_est3:\*
- 13: gb\_est4:\*
- 14: gb\_est5:\*
- 15: em\_estfun:\*
- 16: em\_estom:\*
- 17: em\_gss\_hum:\*
- 18: em\_gss\_inv:\*
- 19: em\_gss\_pln:\*
- 20: em\_gss\_vrt:\*
- 21: em\_gss\_fun:\*
- 22: em\_gss\_mam:\*
- 23: em\_gss\_mus:\*
- 24: em\_gss\_pro:\*
- 25: em\_gss\_rod:\*
- 26: em\_gss\_phg:\*
- 27: em\_gss\_vrl:\*
- 28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match %	Query Length	DB ID	Description
1	145.5	22.3	289	29	U82114 U82114 orde
2	125.5	19.2	1438	28	BH770798 LMGtag54
3	123	18.8	959	29	BZ549048 pacs1-60
4	122.5	18.8	726	14	CB679186 OSJNEf02F
5	119.5	18.3	590	14	CB925602 ABAL.22.F
6	114	17.5	758	13	BQ514888 EST622303
7	106	16.2	603	28	AZ934428 BJ_Ba000
8	104	15.9	765	9	AJ558965 AJ558965
9	100.5	15.4	1033	29	BZ561390 pacs1-164
10	100	15.3	947	29	BZ549047 pacs1-60
11	99	15.2	726	28	AQ989479 Rfc00025
12	98	15.0	1093	13	BU557763 AGENCOURT
13	96.5	14.8	790	29	BZ577987 msh2.5667
14	96	14.7	546	9	AW285510 LGI_241_E
15	96	14.7	551	9	AW285527 LGI_241_G
16	95.5	14.6	677	12	BI378928 BFLG1_000
17	95	14.5	543	12	BI721127 1031054B1
18	95	14.5	673	12	BI956121 HVSMEm002
19	94.5	14.5	1127	29	BZ560689 pacs2-164
20	93.5	14.3	885	29	BZ567978 pacs2-164
21	92	14.1	703	14	CD383492 PTMM08663
22	91.5	14.0	600	10	BG143271 Ia95h02.Y
23	91	13.9	701	14	CD204316 HSL_7.H09
24	91	13.9	1098	13	BU553475 AGENCOURT
25	90.5	13.9	728	12	BJ285991 BJ285991
26	89	13.6	509	10	BE195909 HVSMEm009
27	88	13.5	655	14	CB921057 VVD048E11
28	88	13.5	999	12	BM051288 603634270
29	87.5	13.4	600	29	CC345457 OQAF707H
30	87.5	13.4	655	9	AL883551 AL883551
31	87.5	13.4	719	28	AZ933900 BJ_Ba000
32	87.5	13.4	898	10	BF309796 60182449
33	87	13.3	510	13	BQ977384 QH124M07
34	87	13.3	642	12	BJ283688 BJ283688
35	87	13.3	643	12	BJ254998 BJ254998
36	87	13.3	649	13	BQ801199 WHE2811.E
37	87	13.3	656	13	BU016293 OHE12L01
38	87	13.3	799	12	BM943807 UI-M-EH0p
39	87	13.3	1373	13	BQ890329 AGENCOURT
40	86.5	13.2	491	10	BE333846 us28a04.Y
41	86.5	13.2	494	10	BF722271 mab16h10.Y
42	86.5	13.2	518	12	BM569886 Ih99c11.Y
43	86.5	13.2	547	9	AI652699
44	86.5	13.2	555	12	BI439305 ic62807.Y
45	86.5	13.2	555	12	BI441548 ic55908.Y

# ALIGNMENTS

RESULT 1

U82114

LOCUS

DEFINITION

U82114

ACCESSION

U82114.1

VERSION

GSS

KEYWORDS

SOURCE

ORGANISM

REFERENCE

U82114 289 bp DNA linear GSS 16-FEB-2001  
U82114 ordered cosmid library Mycobacterium leprae genomic clone  
cosmid L-373; contig 64, genomic survey sequence.

U82114.1 GI:3647212  
Mycobacterium leprae  
Mycobacterium leprae  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Mycobacterium.  
1 (bases 1 to 289)

**AUTHORS** Silbag, F.S., Cho, S.N., Cole, S.T. and Brennan, P.J.  
**TITLE** Characterization of a 34-kilodalton protein of *Mycobacterium leprae* that is isologous to the immunodominant 34-kilodalton antigen of *Mycobacterium paratuberculosis*  
**JOURNAL** Infect. Immun. 66 (11), 5576-5579 (1998)  
**MEDLINE** 99003183  
**PUBMED** 9784577  
**COMMENT** Contact: Silbag FS

Eiglmeier, K., Honore, N., Woods, S.A., Caudron, B. and Cole, S.T. Use of an ordered cosmid library to deduce the genomic organization of *Mycobacterium leprae*. *Mol. Microbiol.* 7 (2), 197-206 (1993)  
Class: unknown.

```

FEATURES
  source
    Location/Qualifiers
      1..289
        /organism="Mycobacterium leprae"
        /mol_type="genomic DNA"
        /db_xref="taxon:1769"
        /clone="cosmid L-373; contig 64"
        /clone_lib="ordered cosmid library"
BASE COUNT
  62 a 83 c 93 g 51 t

```

```

BASE COUNT      62 a      83 c      93 g      51 t
ORIGIN          /chr10e1000000 - ordered cosmids library

Alignment Scores:
Pred. No.:      7,948-06      Length:      289
Score:          145.50      Matches:      35
Percent Similarity: 58.82%      Conservative: 15
Best Local Similarity: 41.18%      Mismatches: 34
Query Match:    22.28%      Indels:      1
DB:             29      Gaps:      1

```

US-09-684-215A-18 (1-128) x U82114 (1-289)

QY	41	AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsn---ClyAlaargValGlnArg	59
Db	7	CGGTGGCTGGGTGTACAGGTGCACCAGAAAGGCCACCCGGCGCCCAAAAGTTATGGAC	66
QY	60	ValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVallIeThrAla	79
Db	67	GTCGTGGCCGGTGGTGCGCGCGAATGCCCGCTCCCNAAGGGGTCTTCTCACAAG	126
QY	80	ValaspGlyAlaProIleAsnSerAlaThrAlaMetalaAspAlaasnGlyHlsHls	99
Db	127	ETCGACGACCCCTGTATCATGTAGCGCGACGCGTGTGCTGCCGTGGGTCCAAGGCA	186
QY	100	ProGlyAspValIleSerValThrTrpGlnThrLysaserGlyGlyThrArgThrGlyAsn	119
Db	187	CCCGGTGACAAAGTGTGCTGTACCTATCAGAGTACAGTCTGGTAGCAGTCCGACGGTTCAG	246
QY	120	ValThrLeuAlaGlu	124
Db	247	GTACACATCGGCAAG	261

RESULT 2					
BH770798/c					
LOCUS	BH770798	1438 bp	DNA	linear	GSS 01-MAY-2002
DEFINITION	LLMtag541 MG1363 Random Sequence Tag Library Lactococcus lactis subsp. cremoris genomic, genomic survey sequence.				

ACCESSION	BH770798
VERSION	BH770798.1
	GI:20373755

RECORDS	SOURCE	ORGANISM	REFERENCE
555		<i>Lactococcus lactis</i> subsp. <i>cremoris</i>	1 (bases 1 to 1438)
		<i>Lactococcus lactis</i> subsp. <i>cremoris</i>	Bolotin, A., Ehrlich, S.D. and Sorokin, A.
		Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.	Studies of genomes of dairy bacteria <i>Lactococcus lactis</i>
			Sci. Aliments, (2002) in press
			Contact: Sorokin A
			Genetique Microbienne

INRA

CRJ INRA, Domaine de vilvert, 78352 Jouy en Josas cedex, France  
Tel: 33 1 34 65 25 16  
Fax: 33 1 34 65 25 21  
Email: sorckine@jouy.inra.fr  
best homologue in strain IL1403 is htra (95%)  
Class: shotgun  
High quality sequence start: 30  
High quality sequence stop: 1408.

## FEATURES

```

1:taxid=100909
/organism="Lactococcus lactis subsp. cremoris"
/mol_type="genomic DNA"
/strain="MG1363"
/db_xref="taxon:1359"
/clone_lib="MG1363 Random Sequence Tag Library"
/notes="Vector: pSGM2; Site: 1: Smal; Library of
chromosomal fragments of L.lactis strain MG1363
prepared by partial AluI digestion or by sonica-
tion."
438 a -282 c 260 g 458 t
BASE COUNT

```

	Prepared by particular individual		
	438 a	282 c	260 g
BASE COUNT	458 t		
ORIGIN			

Alignment Scores:		
Pred. No.:	0.00723	Length:
Score:	125.50	Matches:
Percent Similarity:	43.28%	Conservative:
Best Local Similarity:	29.85%	Mismatches:
Query Match:	19.22%	Indels:
DB:	28	Gaps:
		5
		1.38

US-09-684-215A-18 (1-128) x BH770798 (1-1438)

13	GlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAla	-----	27
479	GGTTAGGTGTTTGTCTATCCCATCTAATGATGTGTAACATCATTAATAAACTTGAACAT	420	
28	--GlyGlnIleLysLeuProThrValJHIsleGlyProThrAlaPheLeuGlyLeuGly	46	
419	GATGTAAGATTTCCAGTCCTGCC-----TTAGGTATTCTCT	384	
47	ValValAsp-----AsnAsnGlyAsn-----	53	
383	ATGGTTAGCTTATCTCAATTATCAACAATGATAGTTCTCAACTGAAATTACTAGCAGC	324	
54	-----GlyAlaArgValGlnArgValValGlySerAlaProAlaIleSerLeuGly	70	
323	GTAAGTGTGTGAGTGTGTCTACTCTGTCTCAAGCGGGTCTTCTCTGCCACAGCTGTG	264	
71	IleSerThrGlyAspValIleThrAlaValAlaAspGlyAlaProIleAsnSerAlaThrAla	90	
263	CTGAAGTGCAGTGTGAATCAAGGTGGAGATACGCCGCTTACTTCATCAACAGAC	204	
91	MetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerValThrTrpGlnThr	110	
203	TTACAAAGTCTTCTTACTCACACAATATTAAATGATAGTGTGAAAGTCACCTTAC	147	
111	LysSerGlyGlyThrArgThrGlyAsnValThrLeuAlaGlu	124	
146	CGTGATGTAATTCAGCCACAGCAAAATGTCAAATCTCTAAA	105	

### RESULT 3

BZ549048/c

LOCUS

### DEFINITION

2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11

ACCESSION  
VERSION

VERSION  
KEYWORDS

**KEYWORDS**  
**SOURCE**

SOURCE ORGANISM

**ORIGINATOR**

## REFERENCE

**AUTHORS** Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M., Burns,J.L., Kaul,R. and Olsen,M.V.  
**TITLE** Whole-Genome-Sequence variation among multiple isolates of *Pseudomonas aeruginosa* library  
**JOURNAL** J. Bacteriol., (2002) In press  
**COMMENT** Contact: Chris K. Raymond  
 Genome Center  
 University of Washington  
 Box 352145, Seattle, WA 98105-2145, USA  
 Tel: 2062216954  
 Fax: 2066857244  
 Email: craymond@u.washington.edu  
 Class: shotgun.

**FEATURES**  
 source  
 1. 959  
 Location/Qualifiers  
 /organism="Pseudomonas aeruginosa"  
 /mol\_type="genomic DNA"  
 /strain="1-60"  
 /db\_xref="taxon:287"  
 /clone="pacs1-60\_1644"  
 /clone\_lib="pacs1-60"  
 /note="clinical isolate 1-60 Whole genomic shotgun library."

BASE COUNT 157 a 314 c 291 g 195 t 2 others  
 ORIGIN

Alignment Scores:  
 Pred. No.: 0.00791 Length: 959  
 Score: 123.00 Matches: 41  
 Percent Similarity: 44.88% Conservative: 16  
 Best Local Similarity: 32.28% Mismatches: 46  
 Query Match: 18.84% Indels: 24  
 DB: 29 Gaps: 4

US-09-684-215A-18 (1-128) x BZ549048 (1-959)

Qy 4 SerAspAsnPhGlnLeuSerGlnGly-----GlyGlnGlyPheAlaIleProIleGly 21  
 Db 378 TCGCAGATCTTCACCCGCTCCGGCGCTTCATGGCGCTTCCTTCGCCATTCGCATCGAT 319

Qy 22 GlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThrAla 41  
 Db 318 GTCCGCGTGAACGTCGCGGACCATGTTGAAG-----AAAGCGGCAAGGTGAGT 271

Qy 42 PheLeuGlyLeuGlyValValAlaAspAsnGlyAsn----- 53  
 Db 270 CGCGCGTGGCTGGCGGTGGTGCATCCAGGAAGTGAACAGGATCCGCGGAGTCTCTACGCG 211

Qy 54 -----GlyAlaArgValGlnArgValGlnArgValGlySerAlaProAlaAlaSer 68  
 Db 210 CTGCACAGCCGTCGCGCGCGCTGGTGGCGCAACTGGTGAAGACGCGCGCGGCGCAAG 151

Qy 69 LeuGlyIleSerThrGlyAspValIleThrAlaValAlaAspGlyAlaProIleAsnSerAla 88  
 Db 150 GGTGGCTGCAGGTGGCGGATGATCATAGTTGAACGCCAGTCGATCAACGAGTCC 91

Qy 89 ThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerValThrTrp 108  
 Db 90 GCCGACCTGCGGACCTGGTGGCAACATGAAGCGGCGGACAGATCAACCTG----- 37

Qy 109 GlnThrLysSerGlyGlyThr 115  
 Db 36 -----GACGGGGGATCC 25

RESULT 4  
 CB679186/c 726 bp mRNA linear EST 09-APR-2003  
 LOCUS OSJNEF02F19.r OSJNEF Oryza sativa (japonica cultivar-group) cDNA  
 DEFINITION clone OSJNEF02F19 3', mRNA sequence.  
 ACCESSION CB679186  
 VERSION CB679186.1 GI:29682911  
 KEYWORDS EST.  
 SOURCE Oryza sativa (japonica cultivar-group)

## ORGANISM

Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE  
AUTHORS

1 (bases 1 to 726)  
 Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,  
 Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.

## TITLE

Large-scale identification of ESTs involved in the interaction  
 between rice and Magnaporthe grisea

## JOURNAL

Unpublished

## COMMENT

Contact: Rod Wing  
 Arizona Genomics Institute  
 University of Arizona  
 Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ  
 85721-0088, USA  
 Tel: 520 626 3967  
 Fax: 520 621 9288  
 Email: http://genome.arizona.edu

## PCR Primers

FORWARD: gta aac cga cgg cca gtg  
 BACKWARD: gga aac agc tat gac cat g

## Plate:

02 row: F column: 19

## Seq primer:

gga aac agc tat gac cat g.

## Location/Qualifiers

1. 726  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /mol\_type="mRNA"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:39947"  
 /clone="OSJNEF02F19"  
 /tissue\_type="Leaf"  
 /dev\_stage="3 week"  
 /lab\_host="DH10B"  
 /clone\_lib="OSJNEF"

## note:

Vector: pBluescript II KS +; Site\_1: EcoRI; Site\_2:  
 XhoI; Uninfected Control.

## BASE COUNT

221 a 181 c 151 g 173 t

## ORIGIN

Alignment Scores:  
 Pred. No.: 0.00629 Length: 726  
 Score: 122.50 Matches: 41  
 Percent Similarity: 42.86% Conservative: 16  
 Best Local Similarity: 30.83% Mismatches: 49  
 Query Match: 18.76% Indels: 27  
 DB: 14 Gaps: 4

## US-09-684-215A-18 (1-128) x CB679186 (1-726)

## Qy

10 SerGlnGlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGln 29

## Db

579 ACATCTGCTGCTGTTGTTTGGCCATCCGTCATCAACTGTTCTGAAATAAGCTCCACAG 520

## Qy

30 IleLysLeuProThrValHisIleGlyProThrAlaPheLeuGlyLeuGlyValAlaAsp 49

## Db

519 TTA-----ATTACGTTTGAAAGTCCCGTGTGCTGCTGAATGTGGAGTTT 472

## Qy

50 -----AsnAsnGlyAsnGlyAlaArgValGlnArgVal 60

## Db

471 GCTCCAGATCCATCGCATATACAGTTAATGTTCCGACTGGATCATATAATATTCAGGTT 412

## Qy

61 ValGlySerAlaProAlaAlaSerLeuGly----- 70

## Db

411 CTTGGGGCAGTGTCTGCGAGCAAAAGCTGGTCTTCTTCTACCACTAGGGGTTTGTCTGGT 352

## Qy

71 ---IleSerThrGlyAspValIleThrAlaValAlaAspGlyAlaProIleAsnSerAlaThr 89

## Db

351 ACCATTGCTTGTGTGATGATGTTGTCGCGACCGTAAACATATCAAGGCAAACTCT 292

## Qy

90 AlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerValThrTrpGln 109

## Db

291 GATCTATCGAGGGTTCGTGATGCTTATGCGGCTTGAGCAAGGTGAGCTTGACATCCAA 232

```

QY      110 ThrLysSerGlyGlyThrArgThrGlyAsnValThrLeu 122
Db      231 AGA-----GGCGCTGAATCTCTGGAGGTAACCTTG 202

RESULT 5
CB925602
LOCUS   CB925602
DEFINITION
ABAI_22_F06.b1_A012 Abscissic acid-treated seedlings Sorghum bicolor
cDNA clone ABAI_22_F06_A012 3', mRNA sequence.
ACCESSION
CB925602
VERSION
CB925602.1 GI:30161873
KEYWORDS
EST.
SOURCE
Sorghum bicolor (sorghum)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE
1 (bases 1 to 590)
Cordonnier-Pratt,M.-M., Wentzel,V., Suzuki,Y., Sugano,S., Klein
,R.R., Liang,C., Sun,F., Sullivan,R., Shah,M., Buchanan,C.D.,
Eastman,A. and Pratt,L.H.
An EST database from Sorghum: ABAI-treated seedlings
Unpublished
Other_ESTs: ABAI_22_F06.g1_A012
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in
the Human Genome Center, University of Tokyo Institute of Medical
Science; plant material and RNA prepared at Texas A & M University;
sequencing done in the Laboratory for Genomics and Bioinformatics,
University of Georgia. Sequence ends have been trimmed to exclude
vector and regions below Phred quality 16. Three-prime sequences
are presented as their reverse complement and have been trimmed to
exclude polyA.
Seq primer: Sug3 (CGACCTGCAGCTCGAGCACCA)
POLYA=yes.

FEATURES
Location/Qualifiers
source
1..590
/organism="Sorghum bicolor"
/mol_type="mRNA"
/cultivar="IS3620C"
/db_xref="taxon:4558"
/clone="ABAI_22_F06_A012"
/lab_host="PH10B-r1 phage-resistant E. coli"
/clone_lib="Abscissic acid-treated seedlings"
/notes="Vector: pME18S-FL3; Site_1: XhoI; Site_2: XhoI; The
library was prepared from polyA+ RNA from seedlings grown
in hydroponic culture. After 12 days, medium was
supplemented with 1 mM abscissic acid (ABA), while leaves
were misted with a solution of 1 mM ABA. Roots and leaves
were harvested after 3, 6, 12, and 24 hr and material from
all time points was combined prior to RNA isolation.
Double-stranded cDNA was cloned unidirectionally into
different DraIII sites of the pME18S-FL3 vector (5'-prime
DraIII site is CACTGTGTG, 3'-prime DraIII site is CACCAGTGTG
). XhoI excises the cDNA insert."
BASE COUNT 136 a 137 c 156 g 161 t
ORIGIN

Alignment Scores:
Pred. No.: 0.00995 Length: 590
Score: 119.50 Matches: 38
Percent Similarity: 43.41% Conservative: 18
Best Local Similarity: 29.46% Mismatches: 40
Query Match: 18.30% Indels: 33
DB: 14 Gaps: 5

US-09-684-215A-18 (1-128) x CB925602 (1-590)

QY      15 GlyPheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThr 34
Db      6 GCCTTGTCTATTCCATCATCACTGTACTTAAATCGCTCCTCAGTTA-----53

RESULT 6
BQ514888
LOCUS   BQ514888
DEFINITION
ESTG22303 Generation of a set of potato cDNA clones for microarray
analyses mixed potato tissues Solanum tuberosum cDNA clone STMIO81
3' end, mRNA sequence.
ACCESSION
BQ514888
VERSION
BQ514888.1 GI:21373757
KEYWORDS
EST.
SOURCE
Solanum tuberosum (potato)
ORGANISM
Solanum tuberosum; Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE
1 (bases 1 to 758)
Buell,C.R., Hart,A., Baker,B., Tanksley,S., Fry,W.W., Smart,C.,
Restrepo,S., Griffiths,H., van der Hoeven,R., Tsai,J. and
Karamycheva,S.A.
Generation of a set of potato cDNA clones for microarray analyses
Unpublished
Other_ESTs: ESTG622302
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: T7.

FEATURES
Location/Qualifiers
source
1..758
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec or Binjete"
/db_xref="taxon:4113"
/clone="STMIO81"
/tissue_type="mixed tissues"
/lab_host="SOLR"
/clone_lib="Generation of a set of potato cDNA clones for
microarray analyses mixed potato tissues"
/notes="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Combination of untreated and phytophthora
infestans-treated libraries of stolons, leaves, leaflets,
axillary buds of stem explants, petioles, germinating eyes

```

BASE COUNT 215 a 181 c 128 g 234 t  
ORIGIN

Alignment Scores:  
Pred. NO.: 0.051 Length: 758  
Score: 114.00 Matches: 34  
Percent Similarity: 42.61% Conservative: 15  
Best Local Similarity: 29.57% Mismatches: 42  
Query Match: 17.46% Indels: 24  
DB: 13 Gaps: 3

US-09-684-215A-18 (1-128) x BQ514888 (1-758)

Qy 10 SerGlnGlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGln 29  
Db 592 ACATCAGCAGGTGTGGATTGCAATCCCTTCACACTGTGTGAGAGTGTGCCCCAG 533  
Qy 30 IleLysLeuProThrValHisIleGlyProThrAlaPheLeuGlyLeuGlyValValAsp 49  
Db 532 TTG-----ATCCAATCTGGAAAGTCTTCGTGCTGTTTGAATATTGAAATC 485  
Qy 50 -----AsnAsnGlyAsnGlyAlaArgValGlnArgVal 60  
Db 484 GCTCCAGACCTGATGTCACCAACCACTTAATGTTGCAATGAGCAGTGGTTTCTGGTA 425  
Qy 61 ValGlySerAlaProAlaAlaSerLeuGly-----TCTACTACCGGGTTTTCAGGA 365  
Db 424 CCTGGAATAGTCTCGCAGCGAAAGCGGACTTCTTCTACTACCGGGTTTTCAGGA 365  
Qy 71 ---IleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThr 89  
Db 364 AATATAGTGTGCGGATATTATTGAAGCAGTGTGACAAACCTGTTAGAGTAAAGCA 305  
Qy 90 AlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIle 104  
Db 304 GAGTTGTATAAGCCCTGGTAACTATAACATAGTGTGAAGTT 260

RESULT 7  
AZ934428 603 bp DNA linear GSS 24-APR-2001  
LOCUS  
DEFINITION BJ\_Ba0002I08r B. japonicum BAC library Bradyrhizobium japonicum genomic survey sequence.  
ACCESSION  
VERSION AZ934428.1 GI:13776488  
KEYWORDS  
SOURCE GSS.  
ORGANISM Bradyrhizobium japonicum  
Bacteria; Bradyrhizobium japonicum  
Rhizobiales; Bradyrhizobium japonicum  
Bradyrhizobiaceae; Bradyrhizobium.

REFERENCE  
AUTHORS Tomkins, J.P., Wood, T.C., Stacey, M.G., Loh, J.T., Judd, A., Goicoechea, J.L., Stacey, G., Sadowsky, M.J. and Wing, R.A.  
TITLE A marker-dense, sequence-ready map of the Bradyrhizobium japonicum genome  
JOURNAL Genome Res. 11 (8), 1434-1440 (2001)  
MEDLINE 21376150  
PUBMED 11483585  
COMMENT Contact: Wing RA  
Clemson University Genomics Institute  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 7288  
Fax: 864 656 4293  
Email: rwing@clemson.edu  
Class: BAC ends  
High quality sequence stop: 553.  
Location/Qualifiers  
1. .603  
/organism="Bradyrhizobium japonicum"  
/mol\_type="genomic DNA"  
/strain="USD110"  
/db\_xref="taxon:375"

/lab\_host="E. coli"  
/clone\_lib="B. japonicum BAC library"  
/note="Vector: pIndigo536; Site\_1: HindIII"

BASE COUNT 103 a 201 c 203 g 90 t  
ORIGIN

Alignment Scores:  
Pred. NO.: 0.261 Length: 603  
Score: 106.00 Matches: 35  
Percent Similarity: 41.80% Conservative: 16  
Best Local Similarity: 28.69% Mismatches: 55  
Query Match: 16.23% Indels: 16  
DB: 28 Gaps: 4

US-09-684-215A-18 (1-128) x AZ934428 (1-603)

Qy 13 GlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGlnIleLys--- 31  
Db 245 GGCAATCGCTTCTCGATCCCGGCCAACACCGTGAAGACGGTTGTGCCAGCTCAAGGAC 304  
Qy 32 -----LeuProThrValHisIleGlyProThrAlaPheLeuGly 44  
Db 305 AAGGTTCTGTCAGCCCGCTGGATCGCGCTGCAGATTTCAGCGGTGACG----- 355  
Qy 45 LeuGlyValValAspAsnAsnGly-----AsnGlyAlaArgValGlnArgVal 60  
Db 356 TCGGATATCCCGCAGACGCTCGGCATGAAGAAGCGCGGCGCTGGTGGCGGAGCCG 415  
Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
Db 416 CAGCGAAACGGTCCGCGCGAAGCGCGCATCGAGTCCCGCGAGTGACCTCGCTCGTC 475  
Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100  
Db 476 AACGGCGAATCCGTCAGGAGCGCGCGAGCTCGCCCGCACCATCGCGCGCATCGCCGC 535  
Qy 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
Db 536 CGTGGATCGTGAAGCTTAACGTGCTGCACAAG---GGCCAGGACAATGTCGTGAACCTC 592  
Qy 121 ThrLeu 122  
Db 593 ACCCTC 598

RESULT 8  
AJ558965 765 bp mRNA linear EST 12-JUN-2003  
LOCUS  
DEFINITION AJ558965 Antirrhinum majus whole plant Antirrhinum majus cDNA clone 018\_1\_09\_a08, mRNA sequence.  
ACCESSION  
VERSION AJ558965  
KEYWORDS  
SOURCE EST.  
ORGANISM Antirrhinum majus (snapdragon)  
Antirrhinum majus  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Lamiales; Antirrhinaceae; Antirrhineae;  
Antirrhinum.

REFERENCE  
AUTHORS 1 (bases 1 to 765)  
TITLE Zachgo, S., Stueber, K., Saedler, H., Sommer, H. and Schwarz-Sommer, Z.  
JOURNAL Antirrhinum EST collection  
COMMENT Unpublished  
Contact: Schwarz-Sommer Z  
Molekulare Pflanzen-genetik  
MPI fuer Zuechtungs-forschung  
Carl-von-Linne Weg 10, D-50829, Germany.  
Location/Qualifiers  
1. .765  
/organism="Antirrhinum majus"  
/mol\_type="mRNA"  
/db\_xref="taxon:4151"  
/clone="018\_1\_09\_a08"  
/tissue\_type="whole plant"







BZ577987	BZ577987	790 bp	DNA	linear	GSS 17-DEC-2002
LOCUS	msh2_5667.x1 msh	<i>Pseudomonas aeruginosa</i>	genomic clone	msh2_5667,	
DEFINITION	genomic survey sequence.				
ACCESSION	BZ577987				
VERSION	BZ577987.1	GI:27213048			
KEYWORDS	GSS.				
SOURCE	<i>Pseudomonas aeruginosa</i>				
ORGANISM	<i>Pseudomonas aeruginosa</i>				
	Bacteria: Proteobacteria; Gammaproteobacteria; Pseudomonadales;				
	Pseudomonadaceae; Pseudomonas.				
REFERENCE	1 (bases 1 to 790)				
AUTHORS	Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,				
	Burns,J.L., Kaul,R. and Olsen,M.V.				
TITLE	Whole-Genome Sequence variation among multiple isolates of				
	<i>Pseudomonas aeruginosa</i> library				
JOURNAL	J. Bacteriol., (2002) In press				
COMMENT	Contact: Chris K. Raymond				
	Genome Center				
	University of Washington				
	Box 352145, Seattle, WA 98105-2145, USA				
	Tel: 20622216954				
	Fax: 2066857244				
	Email: craymond@u.washington.edu				
	Class: shotgun.				

```

FEATURES
Location/Qualifiers
1..790
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="MSH"
/db_xref="taxon:287"
/clone_lib="msh_5667"
/clone_lib="msh"
/note="Environmental isolate. Whole genomic shotgun library."
BASE COUNT      107 a   243 c    287 g   150 t     3 others
ORIGIN
Alignment Scores:
Pred. No.:          3.58           Length:         790
Score:             96.50           Matches:         37
Percent Similarity: 42.40%        Conservative:    16
Best Local Similarity: 29.60%      Mismatches:     47
Query Match:       14.78%         Indels:         25
DB:                29            Gaps:           6

US-09-684-215A-18 (1-128) x BZ577987 (1-790)

Qy      9 LeuSerGlnGlyGlyGlnGlyPheAlaIleProIleGly-----GlnAla 23
|||::: |||::: ||| ||| |||
Db      110 CTTGCCGCTCAGGGCAGCACACCACCGCCAGCAGGAGGAGTCTCGTCGTCTTGCC 169

Qy      24 MetAlaIleAlaGlyGlnIleLysLeu---ProThrValHisIleGlyProThrAlaPhe 42
:::||||::: |||:::||||::: ||| ||| ||| :::|||::: |||
Db      170 GTCCCGGTCCAGGGAAACTGNATGTGATCCGGCCGTACTGTGTCGTGCAGATCTCC 229

Qy      43 LeuGlyLeuGlyValValAspAsnAnglyAsnGlyAlaArgValGlnArgValValGly 62
|||||| |::::: ||| ||| ||| |||
Db      230 GCCGCGCTTGCC-----CAACCACGAGGCGGTCTGGGTTCGATCGCGGCGTTGG 280

Qy      63 SerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaValAspGly 82
||| ||| :::::||| ||||| ||| ::| |||||
Db      281 TGTGACCGGGGGGGGTAGCGGTGGCGTTGGGGATCGCCTCGAACGGGT----- 334

Qy      83 AlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisProGlyAsp 102
||| :::: ||| ||| ||| ||| |||
Db      335 GCGGTAGTGGCGTGGCTGGCGTGGGGCGAC-----CGAGGTCAGCACCCAGTCGAT 388

Qy      103 ValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnValThrLeu 122
||| :::: |||
Db      389 ATTCAAGCATCGTC-----GTCAATGCC 412

Qy      123 AlaGluGlyProPro 127

```

```

Db      413 GCTGAGGTGAACCA 427
|||||||
RESULT 14
LOCUS   AW285510                      546 bp  mRNA  linear  EST 19-JUL-2000
DEFINITION LG1_241_E05.g1_A002 Light Grown 1 (LGI) Sorghum bicolor cDNA, mRNA
sequence.
ACCESSION AW285510
VERSION    AW285510.1 GI:6675354
KEYWORDS   EST.
SOURCE     Sorghum bicolor (sorghum)
ORGANISM   Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 546)
Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C. and Pratt,L.H.
An EST database from Sorghum: light-grown seedlings
Unpublished
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions
below Phred quality 16. The threshold for highest quality sequence
is 20.
Seq primer: T7
High quality sequence start: 126
High quality sequence stop: 546
POLYA=Yes.

FEATURES             Location/Qualifiers
     source           1..546
                     /organism="Sorghum bicolor"
                     /mol_type="mRNA"
                     /db_xref="taxon:4558"
                     /clone_lib="Light Grown 1 (LGI)"
                     /note="Organ: 10- to 14-day-old light-grown (greenhouse)
                     seedlings; Vector: Lambda Zap; Site_1: XhoI; Site_2: EcoRI
                     ; The library was made from poly-A RNA in the cloning
                     vector lambda ZAP II. Clones to be sequenced were
                     prepared by mass excision."
BASE COUNT          121 a 127 c 149 g 149 t
ORIGIN
Alignment Scores:
Pred. No.:          2.53          Length:          546
Score:              96.00         Matches:         25
Percent Similarity: 47.62%        Conservative:    15
Best Local Similarity: 29.76%     Mismatches:     26
Query Match:        14.70%       Indels:         18
DB:                  9           Gaps:              2

US-09-684-215A-18 (1-128) x AW285510 (1-546)

Qy      50 AsnAsnGlyAsnGlyAlaArgValGlnArgValValGlySerAlaProAlaAlaSerLeu 69
|||||
Db      56 AATGTTCCGACGAGGAGCTTTACTTAAAGTACCTGGGGGCGAGTCTGCAGCCAAAGCA 115
|||||
Qy      70 GlyIleSer-----
|||||
Db      116 GGTCTTGCTCCACCGCGAGGGTTTGTCTGGTAATATTGTTCTGGGTGATATCATCGTT 175
|||||
Qy      79 AlavalaspGlyAlaProIleAsnSerAlaThralaMetAlaaspAlaLeuAsnGlyHis 98
|||||
Db      176 GCACTGGACGGCAACCTGTTAAGGCGCAATCTGAOCTGTGAGGGTCTCGGATGACTAT 235
|||||
Qy      99 HisProGlyaspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGly 118
|||||
Db      236 GCGCTCGAGATCAGGTGACCTTGACA-----ATCCGCGGAGGC 274
|||||
Qy      119 AsnValThrLeu 122
|||||
Db      275 TCAGAAACCCCTT 286
|||||

RESULT 15
LOCUS   AW285527                      551 bp  mRNA  linear  EST 19-JUL-2000
DEFINITION LG1_241_G05.g1_A002 Light Grown 1 (LGI) Sorghum bicolor cDNA, mRNA
sequence.
ACCESSION AW285527
VERSION    AW285527.1 GI:6675371
KEYWORDS   EST.
SOURCE     Sorghum bicolor (sorghum)
ORGANISM   Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 551)
Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C. and Pratt,L.H.
An EST database from Sorghum: light-grown seedlings
Unpublished
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions
below Phred quality 16. The threshold for highest quality sequence
is 20.
Seq primer: T7
High quality sequence start: 16
High quality sequence stop: 552
POLYA=Yes.

FEATURES             Location/Qualifiers
     source           1..551
                     /organism="Sorghum bicolor"
                     /mol_type="mRNA"
                     /db_xref="taxon:4558"
                     /clone_lib="Light Grown 1 (LGI)"
                     /note="Organ: 10- to 14-day-old light-grown (greenhouse)
                     seedlings; Vector: Lambda Zap; Site_1: XhoI; Site_2: EcoRI
                     ; The library was made from poly-A RNA in the cloning
                     vector lambda ZAP II. Clones to be sequenced were
                     prepared by mass excision."
BASE COUNT          124 a 127 c 150 g 150 t
ORIGIN
Alignment Scores:
Pred. No.:          2.56          Length:          551
Score:              96.00         Matches:         25
Percent Similarity: 47.62%        Conservative:    15
Best Local Similarity: 29.76%     Mismatches:     26
Query Match:        14.70%       Indels:         18
DB:                  9           Gaps:              2

US-09-684-215A-18 (1-128) x AW285527 (1-551)

Qy      50 AsnAsnGlyAsnGlyAlaArgValGlnArgValValGlySerAlaProAlaAlaSerLeu 69
|||||
Db      61 AATGTTCCGACGAGGAGCTTTACTTAAAGTACCTGGGGGCGAGTCTGCAGCCAAAGCA 120
|||||
Qy      70 GlyIleSer-----
|||||
Db      121 GGTCTTGCTCCACCGCGAGGGTTTGTCTGGTAATATTGTTCTGGGTGATATCATCGTT 180
|||||
Qy      79 AlavalaspGlyAlaProIleAsnSerAlaThralaMetAlaaspAlaLeuAsnGlyHis 98
|||||
Db      181 GCAGTGGACGGCAACCTGTTAAGGCGCAATCTGACCTGCTGAGGGTCTCGGATGACTAT 240
|||||
Qy      99 HisProGlyaspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGly 118
|||||

```

```

Db      241 GCGTCGGAGATCAGGTGACCTTGACA-----ATCCGGCGAGGC 279
QY      119 AsnValThrLeu 122
Db      280 TCAGAAACCCCTT 291

```

Search completed: September 5, 2003, 11:54:41  
Job time : 2216.31 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 5, 2003, 08:54:40 ; Search time 185.821 Seconds  
(without alignments)  
1584.308 Million cell updates/sec

Title: US-09-684-215A-18

Perfect score: 653

Sequence: 1 TAASDNFQLSQGGGGAIFI.....QTKSGGTRTGNVTLAEGPPA 128

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1537136 seqs, 1149988732 residues

Total number of hits satisfying chosen parameters: 3074272

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-O=/cgn2\_1/USPTO\_spool\_p/US09684215/runat\_05092003\_072203\_619/app\_query.fasta\_1.853  
-DB=PublishedApplications\_NA -QFWT=fastq -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09684215 -cgn1\_1.130 -runat\_05092003\_072203\_619  
-NCPU=6 -ICPU=3 -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

PublishedApplications\_NA.\*  
1: /cgn2\_1/pubdata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/pubdata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/pubdata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/pubdata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/pubdata/1/pubpna/PCTUS\_NEW\_PUB.seq.\*  
6: /cgn2\_6/pubdata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/pubdata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/pubdata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/pubdata/1/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/pubdata/1/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/pubdata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/pubdata/1/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/pubdata/1/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/pubdata/1/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/pubdata/1/pubpna/US10\_NEW\_PUB.seq.\*  
16: /cgn2\_6/pubdata/1/pubpna/US60\_NEW\_PUB.seq.\*  
17: /cgn2\_6/pubdata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	653	100.0	675	9	US-09-759-143-822 Sequence 822, App

2	653	100.0	675	9	US-09-780-669-822 Sequence 822, App
3	653	100.0	675	9	US-09-822-827-822 Sequence 822, App
4	653	100.0	675	10	US-09-895-793-822 Sequence 822, App
5	653	100.0	675	10	US-09-895-814-822 Sequence 822, App
6	653	100.0	675	12	US-10-144-678A-822 Sequence 822, App
7	653	100.0	675	13	US-10-012-896-822 Sequence 822, App
8	653	100.0	822	10	US-09-736-457-1862 Sequence 1862, App
9	653	100.0	822	10	US-09-902-941-1862 Sequence 1862, App
10	653	100.0	822	10	US-09-849-626-1862 Sequence 1862, App
11	653	100.0	822	14	US-10-017-754-1862 Sequence 1862, App
12	653	100.0	900	9	US-09-735-705-353 Sequence 353, App
13	653	100.0	900	10	US-09-850-716A-353 Sequence 353, App
14	653	100.0	900	10	US-09-897-778-353 Sequence 353, App
15	653	100.0	900	12	US-10-117-982-353 Sequence 353, App
16	653	100.0	915	9	US-09-759-143-834 Sequence 834, App
17	653	100.0	915	9	US-09-780-669-834 Sequence 834, App
18	653	100.0	915	9	US-09-822-827-834 Sequence 834, App
19	653	100.0	915	10	US-09-895-793-834 Sequence 834, App
20	653	100.0	915	10	US-09-895-814-834 Sequence 834, App
21	653	100.0	915	12	US-10-144-678A-834 Sequence 834, App
22	653	100.0	915	13	US-10-012-896-834 Sequence 834, App
23	653	100.0	945	10	US-09-736-457-1861 Sequence 1861, App
24	653	100.0	945	10	US-09-902-941-1861 Sequence 1861, App
25	653	100.0	945	10	US-09-849-626-1861 Sequence 1861, App
26	653	100.0	945	14	US-10-017-754-1861 Sequence 1861, App
27	653	100.0	1012	9	US-09-735-705-351 Sequence 351, App
28	653	100.0	1012	10	US-09-850-716A-351 Sequence 351, App
29	653	100.0	1012	10	US-09-897-778-351 Sequence 351, App
30	653	100.0	1012	12	US-10-117-982-351 Sequence 351, App
31	653	100.0	1035	9	US-09-922-217-1084 Sequence 1084, App
32	653	100.0	1035	10	US-09-833-263-1084 Sequence 1084, App
33	653	100.0	1035	11	US-09-938-864-388 Sequence 388, App
34	653	100.0	1035	13	US-10-025-380-1084 Sequence 1084, App
35	653	100.0	1035	14	US-10-125-635A-388 Sequence 388, App
36	653	100.0	1035	14	US-10-002-603-388 Sequence 388, App
37	653	100.0	1155	10	US-09-902-941-1875 Sequence 1875, App
38	653	100.0	1155	10	US-09-849-626-1875 Sequence 1875, App
39	653	100.0	1155	14	US-10-017-754-1875 Sequence 851, App
40	653	100.0	1203	9	US-09-759-143-851 Sequence 851, App
41	653	100.0	1203	9	US-09-780-669-851 Sequence 851, App
42	653	100.0	1203	9	US-09-822-827-851 Sequence 851, App
43	653	100.0	1203	10	US-09-895-793-851 Sequence 851, App
44	653	100.0	1203	10	US-09-895-814-851 Sequence 851, App
45	653	100.0	1203	12	US-10-144-678A-851 Sequence 851, App

#### ALIGNMENTS

RESULT 1  
US-09-759-143-822  
Sequence 822, Application US/09759143  
Patent No. US200202248A1  
GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun  
APPLICANT: Dillon, Davin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Jiang, Yuqi  
APPLICANT: Henderson, Robert A.  
APPLICANT: Kalos, Michael D.  
APPLICANT: Fanger, Gary R.  
APPLICANT: Retter, Marc W.  
APPLICANT: Stolk, John A.  
APPLICANT: Day, Craig H.  
APPLICANT: Vedrick, Thomas S.  
APPLICANT: Carter, Darrick  
APPLICANT: Li, Samuel  
APPLICANT: Wang, Aijun  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Hepler, William  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
DIAGNOSIS OF PROSTATE CANCER  
FILE REFERENCE: 210121.427C23

Score:

Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 9 Indels: 0  
DB: Gaps: 0

US-09-684-215A-18 (1-128) x US-09-822-827-822 (1-675)

```
Qy 1 ThrAlaAlaSerAspAsnPhGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 22 ACGGCGCGTCCGATACCTCCAGCTGCCAGGGTGGCAGGATTCGCATTCGCGATC 81
Qy 21 GlyClnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
Db 82 GGGCAGCGATGGCGATCGCGGCCAGATCAAGCTCCACCGTTCATATCGGCGCTACC 141
Qy 41 AlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArgValGlnArgVal 60
Db 142 GCCTTCCTCGGCTGGGTGTGTGCGACAACAGCGGCGGCGGCGGCGGCGGCGG 201
Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
Db 202 GTGCGGAGCGCTCGGCGGCAAGTCTCGGCATCTCCACCGGCGAGGTGATCACCGCGGTC 261
Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100
Db 262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGCGGCGAGCGCTTAACGGGCGATCATCCC 321
Qy 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
Db 322 GGTGACGTCTATCTCGGTGACCTGGCAACCAAGTCGGCGGCGGCGGCGGCGGCGG 381
```

## RESULT 4

US-09-895-793-822

; Sequence 822, Application US/09895793

; Publication No. US20020192763A1

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; APPLICANT: Dillon, Davin C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Harlocker, Susan L.

; APPLICANT: Jiang, Yuqiu

; APPLICANT: Kalos, Michael D.

; APPLICANT: Retter, Marc W.

; APPLICANT: Stolk, John A.

; APPLICANT: Day, Craig H.

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Carter, Darrick

; APPLICANT: Li, Samuel X.

; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Hepler, William T.

; APPLICANT: Henderson, Robert A.

; APPLICANT: Hural, John

; APPLICANT: McNeill, Patricia D.

; APPLICANT: Houghton, Raymond L.

; APPLICANT: Vinals de Bassols, Carlota

; APPLICANT: Foy, Teresa

; APPLICANT: Fanger, Gary R.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; FILE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER

; FILE REFERENCE: 210121.534C2

; CURRENT APPLICATION NUMBER: US/09/895,793

; CURRENT FILING DATE: 2001-06-29

; NUMBER OF SEQ ID NOS: 982

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 822

; LENGTH: 675

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-895-793-822

Alignment Scores:

Pred. No.: 7,59e-71 Length: 675

Score: 653.00 Matches: 128

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 10 Gaps: 0

US-09-684-215A-18 (1-128) x US-09-895-793-822 (1-675)

```
Qy 1 ThrAlaAlaSerAspAsnPhGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 22 ACGGCGCGTCCGATACCTCCAGCTGCCAGGGTGGCAGGATTCGCATTCGCGATC 81
Qy 21 GlyClnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
Db 82 GGGCAGCGATGGCGATCGCGGCCAGATCAAGCTCCACCGTTCATATCGGCGCTACC 141
Qy 41 AlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArgValGlnArgVal 60
Db 142 GCCTTCCTCGGCTGGGTGTGTGCGACAACAGCGGCGGCGGCGGCGGCGGCGG 201
Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
Db 202 GTGCGGAGCGCTCGGCGGCAAGTCTCGGCATCTCCACCGGCGAGGTGATCACCGCGGTC 261
Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100
Db 262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGCGGCGAGCGCTTAACGGGCGATCATCCC 321
Qy 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
Db 322 GGTGACGTCTATCTCGGTGACCTGGCAACCAAGTCGGCGGCGGCGGCGGCGGCGG 381
```

## RESULT 5

US-09-895-814-822

; Sequence 822, Application US/09895814

; Publication No. US20020193296A1

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; APPLICANT: Dillon, Davin C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Harlocker, Susan L.

; APPLICANT: Jiang, Yuqiu

; APPLICANT: Kalos, Michael D.

; APPLICANT: Retter, Marc W.

; APPLICANT: Stolk, John A.

; APPLICANT: Day, Craig H.

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Carter, Darrick

; APPLICANT: Li, Samuel X.

; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Hepler, William T.

; APPLICANT: Henderson, Robert A.

; APPLICANT: Hural, John

; APPLICANT: McNeill, Patricia D.

; APPLICANT: Houghton, Raymond L.

; APPLICANT: Vinals de Bassols, Carlota

; APPLICANT: Foy, Teresa

; APPLICANT: Fanger, Gary R.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; FILE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER

; FILE REFERENCE: 210121.427C26

; CURRENT APPLICATION NUMBER: US/09/895,814

; CURRENT FILING DATE: 2001-06-29

; NUMBER OF SEQ ID NOS: 990

```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 822
; LENGTH: 675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-895-814-822

Alignment Scores:
Pred. No.: 7-59e-71 Length: 675
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-684-215A-18 (1-128) x US-09-895-814-822 (1-675)
QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlyGlnGlyPheAlaIlePro 20
DB 22 ACGGCGCGCTCCGATAAATTCAGCTGTCCAGGTGGCAGGATTCGCCATTCGGATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
DB 82 GGGCAGGCGATGGCGATCCGGCCAGATCAAGCTTCCACCGCTCATATCGGGCTACC 141
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
DB 142 GCCTTCCTCGGCTGGGTTCGACAAACACGACGCGGACGAGTCCCAACGCGGTG 201
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
DB 202 GTCCGAGCGCTCCCGCGCAAGTCTCGGCATCTCCACGGCGAGTATCACCGCGGTG 261
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100
DB 262 GACGGCGCTCCGATCAACTCGCCACCGGATGGCGGACGCGCTTAACGGGATCATCC 321
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
DB 322 GGTGACGTATCTCGGTGACCTGGCAACCAAGTCCGGCGGACGCGTACAGGGAACGTG 381
QY 121 ThrLeuAlaGluGlyProProAla 128
DB 382 ACATTGGCGGAGGACCCCGGCC 405

RESULT 6
US-10-144-678A-822
; Sequence 822, Application US/10144678A
; Publication No. US20030157089A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Henderson, Robert A.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aljun
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals y de Bassols, Carlota
; APPLICANT: Foy, Teresa M.
; APPLICANT: Watanabe, Yoshihiro

; APPLICANT: Deng, Ta
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121, 427C28
; CURRENT APPLICATION NUMBER: US/10/144, 678A
; NUMBER OF SEQ ID NOS: 11033
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 822
; LENGTH: 675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-144-678A-822

Alignment Scores:
Pred. No.: 7-59e-71 Length: 675
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

US-09-684-215A-18 (1-128) x US-10-144-678A-822 (1-675)
QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlyGlnGlyPheAlaIlePro 20
DB 22 ACGGCGCGCTCCGATAAATTCAGCTGTCCAGGTGGCAGGATTCGCCATTCGGATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
DB 82 GGGCAGGCGATGGCGATCCGGCCAGATCAAGCTTCCACCGCTCATATCGGGCTACC 141
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
DB 142 GCCTTCCTCGGCTGGGTTCGACAAACACGACGCGGACGAGTCCCAACGCGGTG 201
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
DB 202 GTCCGAGCGCTCCCGCGCAAGTCTCGGCATCTCCACGGCGAGTATCACCGCGGTG 261
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100
DB 262 GACGGCGCTCCGATCAACTCGCCACCGGATGGCGGACGCGCTTAACGGGATCATCC 321
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
DB 322 GGTGACGTATCTCGGTGACCTGGCAACCAAGTCCGGCGGACGCGTACAGGGAACGTG 381
QY 121 ThrLeuAlaGluGlyProProAla 128
DB 382 ACATTGGCGGAGGACCCCGGCC 405

RESULT 7
US-10-012-896-822
; Sequence 822, Application US/10012896
; Publication No. US20020183251A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aljun
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
```



```
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; APPLICANT: Wantanabe, Yoshihiro
; APPLICANT: Meagher, Madeleine Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C27
; CURRENT APPLICATION NUMBER: US/10/012.896
; CURRENT FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 1011
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 822
; LENGTH: 675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-012-896-822

Alignment Scores:
Pred. No.: 7,59e-71 Length: 675
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13

US-09-684-215A-18 (1-128) x US-10-012-896-822 (1-675)

QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 22 ACGGCGGCTCCGATAACTTCCAGCTGTCCAGGGTGGCAGGATTCCCGATCCGATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyClnIleLysLeuProThrValHisIleGlyProThr 40
Db 82 GGGCAGGCGATGGCGATCGCGGGCAGATCAAGCTTCCACCGTTTCATATCGGGCTACC 141
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnGlnGlyGlnGlyPheAlaIleArgVal 60
Db 142 GCCTTCCTCGCTGGTGGTGTGCGACAAACACGCGCAGCGAGTCCACGCGTG 201
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
Db 202 GTCGGGAGCGTCCGCGGCAAGTCTCGGCATCTCCACCGCGAGTGTATCACCGCGTGC 261
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100
Db 262 GACGGCGCTCGATCACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCC 321
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
Db 322 GGTGACGTCATCTCGGTGACCTGGCAACCAAGTGGCGGCGGCACGCGTACAGGACGTCG 381

RESULT 8
US-09-736-457-1862
; Sequence 1862, Application US/09736457
; Patent No. US20020168637A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedrick, Tom
; APPLICANT: Carter, Darriack
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Lique
```

```
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736.457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1862
; LENGTH: 822
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-736-457-1862

Alignment Scores:
Pred. No.: 9,73e-71 Length: 822
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10

US-09-684-215A-18 (1-128) x US-09-736-457-1862 (1-822)

QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 22 ACGGCGGCTCCGATAACTTCCAGCTGTCCAGGGTGGCAGGATTCCCGATCCGATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyClnIleLysLeuProThrValHisIleGlyProThr 40
Db 82 GGGCAGGCGATGGCGATCGCGGGCAGATCAAGCTTCCACCGTTTCATATCGGGCTACC 141
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnGlnGlyGlnGlyPheAlaIleArgVal 60
Db 142 GCCTTCCTCGCTGGTGGTGTGCGACAAACACGCGCAGCGAGTCCACGCGTG 201
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
Db 202 GTCGGGAGCGTCCGCGGCAAGTCTCGGCATCTCCACCGCGAGTGTATCACCGCGTGC 261
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100
Db 262 GACGGCGCTCGATCACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCC 321
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
Db 322 GGTGACGTCATCTCGGTGACCTGGCAACCAAGTGGCGGCGGCACGCGTACAGGACGTCG 381

RESULT 9
US-09-902-941-1862
; Sequence 1862, Application US/09902941
; Patent No. US2002017952A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Wantanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darriack
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedrick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.478C17
; CURRENT APPLICATION NUMBER: US/09/902.941
; CURRENT FILING DATE: 2001-07-10
```

```
; NUMBER OF SEQ ID NOS: 2002
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1862
; LENGTH: 822
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-902-941-1862

Alignment Scores:
Pred. No.: 9,73e-71 Length: 822
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-684-215A-18 (1-128) x US-09-902-941-1862 (1-822)
QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProfile 20
Db 22 ACGGCGCGCTCCGATAACTTCCAGCTGTCCAGGTGGCGAGGATTCGCCATTCCGATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
Db 82 GGGCAGGCGATGGCGATCGCGGCCAGATCAAGCTTCCACCGCTTCATATCGGGCTACC 141
QY 41 AlaPheLeuGlyLeuGlyValValAlaAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
Db 142 GCCTTCTCGGCTTGGGTGTTCTGCACAAACGACGCGGCGAGTCCCAACGCGTG 201
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
Db 202 GTCGGAGCGCTCCGGCGCAAGTCTCGCATCTCCACCGCGGCGATGATCACCAGGTC 261
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100
Db 262 GACGCGCTCCGATCAACTCGGCACCGCATGTCGCGATCTCCACCGCGGCGATCATCC 321
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyThrArgThrGlyAsnVal 120
Db 322 GGTGAGCTCATCTCGTGACCTGGCAACCAAGTCCGGCGGCGACGCTTACAGGAAACG 381
QY 121 ThrLeuAlaGluGlyProProAla 128
Db 382 ACATTGGCGAGGACCCCGGCC 405

RESULT 10
US-09-849-626-1862
; Sequence 1862, Application US/09849626
; Publication No. US2002019769A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya
; APPLICANT: Fanger, Gary
; APPLICANT: Wang, Aljun
; APPLICANT: Wang, Tonglong
; APPLICANT: Switzer, Anne
; APPLICANT: McNeill, Patricia
; APPLICANT: Clapper, Jonathan
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C16
; CURRENT APPLICATION NUMBER: US/09/849,626
; CURRENT FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 1926
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1862
; LENGTH: 822
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-849-626-1862

Alignment Scores:
Pred. No.: 9,73e-71 Length: 822
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-09-684-215A-18 (1-128) x US-09-849-626-1862 (1-822)
QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProfile 20
Db 22 ACGGCGCGCTCCGATAACTTCCAGCTGTCCAGGTGGCGAGGATTCGCCATTCCGATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
Db 82 GGGCAGGCGATGGCGATCGCGGCCAGATCAAGCTTCCACCGCTTCATATCGGGCTACC 141
QY 41 AlaPheLeuGlyLeuGlyValValAlaAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
Db 142 GCCTTCTCGGCTTGGGTGTTCTGCACAAACGACGCGGCGAGTCCCAACGCGTG 201
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
Db 202 GTCGGAGCGCTCCGGCGCAAGTCTCGCATCTCCACCGCGGCGATGATCACCAGGTC 261
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100
Db 262 GACGCGCTCCGATCAACTCGGCACCGCATGTCGCGATCTCCACCGCGGCGATCATCC 321
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyThrArgThrGlyAsnVal 120
Db 322 GGTGAGCTCATCTCGTGACCTGGCAACCAAGTCCGGCGGCGACGCTTACAGGAAACG 381
QY 121 ThrLeuAlaGluGlyProProAla 128
Db 382 ACATTGGCGAGGACCCCGGCC 405

RESULT 11
US-10-017-754-1862
; Sequence 1862, Application US/10017754
; Publication No. US20030054363A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tonglong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.478C18
; CURRENT APPLICATION NUMBER: US/10/017,754
; CURRENT FILING DATE: 2001-10-29
; NUMBER OF SEQ ID NOS: 2004
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1862
; LENGTH: 822
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-017-754-1862

Alignment Scores:
Pred. No.: 9,73e-71 Length: 822
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0
```

US-09-684-215A-18 (1-128) x US-10-017-754-1862 (1-822)

```

Oy 1 ThrAlaAlaSerAspAsnPhleGlnLeuSerGlnGlyGlnGlyPheAlaIleProle 20
Db 22 ACGCGCGGTCGGAATACTTCCAGCTGTCCAGAGGTGGCAGGATTCGCATTCGCATC 81
Oy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
Db 82 GGCAGCGGATGGCGATCGCGGGCCAGATCAAGCTTCCACCGTTTCATATCGGCGCTACC 141
Oy 41 AlaphelLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
Db 142 GCCTTCCTCGGCTTGGGTGTGTGCGACAAACACGCGCAGCGGATTCACCGCGGTG 201
Oy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
Db 142 GCCTTCCTCGGCTTGGGTGTGTGCGACAAACACGCGCAGCGGATTCACCGCGGTG 201
Oy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
Db 202 GTCGGGAGCGCTCGCGGGCAAGTCTCGGCATCTCCACCGCGGATTCACCGCGGTG 261
Oy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
Db 262 GACGGCGCTCGATCAACTCGGCGCGGATCGGCGGATCGGCGGATTCATCCATCC 321
Oy 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
Db 322 GGTGACGTCTCTCGGTGACCTGGCAACCAACGAGTCCGGCGGCGGATTCACCGGTG 381
Oy 121 ThrLeuAlaGluGlyProAla 128
Db 382 ACATTGGCGGAGGACCCCGGCC 405

```

RESULT 12

US-09-735-705-353

; Sequence 353, Application US/09735705

; Patent No. US20020052329A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong

; APPLICANT: Fan, Lijun

; APPLICANT: Kalos, Michael D.

; APPLICANT: Bangur, Chaityanya S.

; APPLICANT: Hosken, Nancy

; APPLICANT: Fanger, Gary R.

; APPLICANT: Li, Samuel X.

; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Henderson, Robert A.

; APPLICANT: McNeill, Patricia D.

; APPLICANT: Fanger, Neil

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE REFERENCE: 210121.455C14

; CURRENT APPLICATION NUMBER: US/09/735,705

; CURRENT FILING DATE: 2000-12-12

; NUMBER OF SEQ ID NOS: 419

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 353

; LENGTH: 900

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-735-705-353

Alignment Scores:

```

Pred. No.: 1.09e-70 Length: 900
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

```

US-09-684-215A-18 (1-128) x US-09-735-705-353 (1-900)

```

Oy 1 ThrAlaAlaSerAspAsnPhleGlnLeuSerGlnGlyGlnGlyPheAlaIleProle 20
Db 22 ACGCGCGGTCGGAATACTTCCAGCTGTCCAGGTTGGCAGGATTCGCATTCGCATC 81

```

```

Oy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
Db 82 GGCAGCGGATGGCGATCGCGGGCCAGATCAAGCTTCCACCGTTTCATATCGGCGCTACC 141
Oy 41 AlaphelLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
Db 142 GCCTTCCTCGGCTTGGGTGTGTGCGACAAACACGCGCAGCGGATTCACCGCGGTG 201
Oy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
Db 202 GTCGGGAGCGCTCGCGGGCAAGTCTCGGCATCTCCACCGCGGATTCACCGCGGTG 261
Oy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
Db 262 GACGGCGCTCGATCAACTCGGCGCGGATCGGCGGATCGGCGGATTCATCCATCC 321
Oy 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
Db 322 GGTGACGTCTCTCGGTGACCTGGCAACCAACGAGTCCGGCGGCGGATTCACCGGTG 381
Oy 121 ThrLeuAlaGluGlyProAla 128
Db 382 ACATTGGCGGAGGACCCCGGCC 405

```

RESULT 13

US-09-850-716A-353

; Sequence 353, Application US/09850716A

; Patent No. US20020115139A1

; GENERAL INFORMATION:

; APPLICANT: Kalos, Michael D.

; APPLICANT: McNeill, Patricia D.

; APPLICANT: Retter, Marc W.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE REFERENCE: 210121.455C15

; CURRENT APPLICATION NUMBER: US/09/850,716A

; CURRENT FILING DATE: 2001-05-07

; NUMBER OF SEQ ID NOS: 440

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 353

; LENGTH: 900

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-850-716A-353

Alignment Scores:

```

Pred. No.: 1.09e-70 Length: 900
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

```

US-09-684-215A-18 (1-128) x US-09-850-716A-353 (1-900)

```

Oy 1 ThrAlaAlaSerAspAsnPhleGlnLeuSerGlnGlyGlnGlyPheAlaIleProle 20
Db 22 ACGCGCGGTCGGAATACTTCCAGCTGTCCAGGTTGGCAGGATTCGCATTCGCATC 81
Oy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
Db 82 GGCAGCGGATGGCGATCGCGGGCCAGATCAAGCTTCCACCGTTTCATATCGGCGCTACC 141
Oy 41 AlaphelLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
Db 142 GCCTTCCTCGGCTTGGGTGTGTGCGACAAACACGCGCAGCGGATTCACCGCGGTG 201
Oy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
Db 202 GTCGGGAGCGCTCGCGGGCAAGTCTCGGCATCTCCACCGCGGATTCACCGCGGTG 261
Oy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100

```

Db 262 GACGGCGTCCGATCAACTCGGCCACCGGATCGCGGACGCGCTTAACGGGATCATCCC 321  
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
Db 322 GGTGACGTATCTCGGTGACCTGGGCAACCAAGTGGGGCGGCACGGGTACAGGGAACGTG 381  
QY 121 ThrLeuAlaGluGlyProProLa 128  
Db 382 ACATTGGCGAGGAGGCCCGGCC 405

## RESULT 14

US-09-897-778-353  
; Sequence 353, Application US/0989778  
; Patent No. US20020147143A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Warnerakis, Margarita  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Vedwick, Thomas S.  
; APPLICANT: Carter, Darrick  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Peckham, David W.  
; APPLICANT: Fanger, Neil  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.455C16  
; CURRENT APPLICATION NUMBER: US/09/897,778  
; NUMBER OF SEQ ID NOS: 467  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 353  
; LENGTH: 900  
; TYPE: DNA  
; ORGANISM: Homo sapiens

## US-09-897-778-353

Alignment Scores:  
Pred. No.: 1.09e-70 Length: 900  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0

US-09-684-215A-18 (1-128) x US-09-897-778-353 (1-900)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20  
Db 22 ACGGGCGCTCCGATAACTTCCAGCTGTCCAGGTGGCGAGGATTCGCCATTCGGATC 81  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
Db 82 GGGCAGCGATCGGATCGCGGCCAGATCAAGCTTCCACCGCTCATATCGGGCTACC 141  
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60  
Db 142 GCCTTCTCGGTGGTGTTCGACAAACACGACGCGGCGAGTCCCAACGCGGTG 201  
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
Db 202 GTCCGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACGGCGAGTGTATCACCGGGTTC 261  
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100  
Db 262 GACGGCGTCCGATCAACTCGCGGATGGCGGAGCGCTTAACGGGATCATCCC 321  
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
Db 322 GGTGACGTATCTCGGTGACCTGGGCAACCAAGTGGGGCGGCACGGGTACAGGGAACGTG 381  
QY 121 ThrLeuAlaGluGlyProProLa 128

Search completed: September 5, 2003, 12:04:47  
Job time: 189.154 secs

Db 382 ACATTGGCGAGGAGGCCCGGCC 405

## RESULT 15

US-10-117-982-353  
; Sequence 353, Application US/10117982  
; Publication No. US20030138438A1  
; GENERAL INFORMATION:  
; APPLICANT: Foy, Teresa M.  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Vedwick, Thomas S.  
; APPLICANT: Carter, Darrick  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Mericle, Barbara  
; APPLICANT: Spies, Gregory A.  
; APPLICANT: Fan, Liqun  
; APPLICANT: Wang, Tongtong  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210120.455C18  
; CURRENT APPLICATION NUMBER: US/10/117,982  
; CURRENT FILING DATE: 2002-04-05  
; NUMBER OF SEQ ID NOS: 484  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 353  
; LENGTH: 900  
; TYPE: DNA  
; ORGANISM: Homo sapiens

## US-10-117-982-353

Alignment Scores:  
Pred. No.: 1.09e-70 Length: 900  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 12 Gaps: 0

US-09-684-215A-18 (1-128) x US-10-117-982-353 (1-900)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20  
Db 22 ACGGGCGCTCCGATAACTTCCAGCTGTCCAGGTGGCGAGGATTCGCCATTCGGATC 81  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
Db 82 GGGCAGCGATCGGATCGCGGCCAGATCAAGCTTCCACCGCTCATATCGGGCTACC 141  
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60  
Db 142 GCCTTCTCGGTGGTGTTCGACAAACACGACGCGGCGAGTCCCAACGCGGTG 201  
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
Db 202 GTCCGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACGGCGAGTGTATCACCGGGTTC 261  
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100  
Db 262 GACGGCGTCCGATCAACTCGCGGATGGCGGAGCGCTTAACGGGATCATCCC 321  
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
Db 322 GGTGACGTATCTCGGTGACCTGGGCAACCAAGTGGGGCGGCACGGGTACAGGGAACGTG 381  
QY 121 ThrLeuAlaGluGlyProProLa 128  
Db 382 ACATTGGCGAGGAGGCCCGGCC 405

Search completed: September 5, 2003, 12:04:47  
Job time: 189.154 secs

---

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 5, 2003, 08:54:40 ; Search time 43.5517 Seconds

(without alignments)  
1584.308 Million cell updates/sec

Title: US-09-684-215A-17

Perfect score: 148

Sequence: 1 TRASDNFQLSQGGGFAIPGQAWAIAGQI 30

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1537136 seqs, 1149988732 residues

Total number of hits satisfying chosen parameters: 3074272

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-Q/cgn2\_1/USPTO.spool.p/US09684215/runat\_05092003\_072203\_619/app.query.fasta\_1.853  
-DB=Published Applications\_NA -QFWT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=BLOSUM62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US09684215 -CGN\_1\_130 @runat\_05092003\_072203\_619  
-NCPU=6 -ICPU=3 -NO\_MMAP -LARGESQRY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications\_NA:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
16: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
17: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUBCOMB.seq:\*  
17: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query Match	Length	ID	Description
1	148	100.0	186	10 US-09-902-941-1879

2	148	100.0	186	10 US-09-849-626-1879
3	148	100.0	186	14 US-10-017-754-1879
4	148	100.0	399	9 US-09-757-417-46
5	148	100.0	399	14 US-10-042-945-46
6	148	100.0	447	12 US-10-084-843-4
7	148	100.0	447	12 US-10-193-002-4
8	148	100.0	675	9 US-09-759-143-822
9	148	100.0	675	9 US-09-780-669-822
10	148	100.0	675	9 US-09-822-827-822
11	148	100.0	675	10 US-09-895-793-822
12	148	100.0	675	10 US-09-895-814-822
13	148	100.0	675	12 US-10-144-678A-822
14	148	100.0	675	13 US-10-012-896-822
15	148	100.0	702	9 US-09-287-849-27
16	148	100.0	702	12 US-10-359-460-27
17	148	100.0	822	10 US-09-736-457-1862
18	148	100.0	822	10 US-09-902-941-1862
19	148	100.0	822	10 US-09-849-626-1862
20	148	100.0	822	14 US-10-017-754-1862
21	148	100.0	861	10 US-09-902-941-1877
22	148	100.0	861	10 US-09-849-626-1877
23	148	100.0	861	14 US-10-017-754-1877
24	148	100.0	900	9 US-09-735-705-353
25	148	100.0	900	10 US-09-850-716A-353
26	148	100.0	900	10 US-09-897-778-353
27	148	100.0	915	12 US-10-117-982-353
28	148	100.0	915	9 US-09-759-143-834
29	148	100.0	915	9 US-09-780-669-834
30	148	100.0	915	9 US-09-822-827-834
31	148	100.0	915	10 US-09-895-793-834
32	148	100.0	915	10 US-09-895-814-834
33	148	100.0	915	12 US-10-144-678A-834
34	148	100.0	915	13 US-10-012-896-834
35	148	100.0	945	10 US-09-736-457-1861
36	148	100.0	945	10 US-09-902-941-1861
37	148	100.0	945	10 US-09-849-626-1861
38	148	100.0	945	14 US-10-017-754-1861
39	148	100.0	1012	9 US-09-735-705-351
40	148	100.0	1012	10 US-09-850-716A-351
41	148	100.0	1012	10 US-09-897-778-351
42	148	100.0	1012	12 US-10-117-982-351
43	148	100.0	1035	9 US-09-922-217-1084
44	148	100.0	1035	10 US-09-833-263-1084
45	148	100.0	1035	11 US-09-938-864-388

#### ALIGNMENTS

RESULT 1  
US-09-902-941-1879  
; Sequence 1879, Application US/09902941  
; Patent No. US20020172952A1  
; GENERAL INFORMATION:  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Johnson, Jeffrey C.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Marnetakis, Margarita  
; APPLICANT: Carter, Darrick  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Vadivick, Thomas S.  
; APPLICANT: McNabb, Andria  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.478C17  
; CURRENT APPLICATION NUMBER: US/09/902,941  
; NUMBER OF SEQ ID NOS: 2002  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1879  
; LENGTH: 186

```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-902-941-1879

Alignment Scores:
Pred. No.:      1,01e-16      Length:      186
Score:          148.00      Matches:      30
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:      0
DB:              10          Gaps:      0

US-09-684-215A-17 (1-30) x US-09-902-941-1879 (1-186)
QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20
Db 22 ACGGCGCGCTCGGATAACTTCCAGCTGTCCAGGTGGCGAGGATTCGCCATTCCGATC 81

QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
Db 82 GGCAGGCGATGGCGATCGCGGCCAGATC 111

RESULT 2
US-09-849-626-1879
; Sequence 1879, Application US/09849626
; Publication No. US20020197669A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya
; APPLICANT: Fanger, Gary
; APPLICANT: Wang, Aljun
; APPLICANT: Wang, Tongtong
; APPLICANT: Switzer, Anne
; APPLICANT: McNeill, Patricia
; APPLICANT: Clapper, Jonathan
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C16
; CURRENT APPLICATION NUMBER: US/09/849,626
; NUMBER OF SEQ ID NOS: 1926
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1879
; LENGTH: 186
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-849-626-1879

Alignment Scores:
Pred. No.:      1,01e-16      Length:      186
Score:          148.00      Matches:      30
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:      0
DB:              10          Gaps:      0

US-09-684-215A-17 (1-30) x US-09-849-626-1879 (1-186)
QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20
Db 22 ACGGCGCGCTCGGATAACTTCCAGCTGTCCAGGTGGCGAGGATTCGCCATTCCGATC 81

QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
Db 82 GGCAGGCGATGGCGATCGCGGCCAGATC 111

RESULT 3
US-10-017-754-1879
; Sequence 1879, Application US/10017754
; Publication No. US20030054363A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-017-754-1879

Alignment Scores:
Pred. No.:      2,44e-16      Length:      399
Score:          148.00      Matches:      30
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:      0
DB:              9          Gaps:      0

US-09-684-215A-17 (1-30) x US-09-757-417-46 (1-399)
QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20
Db 82 GGCAGGCGATGGCGATCGCGGCCAGATC 111

RESULT 4
US-09-757-417-46
; Sequence 46, Application US/09757417
; Patent No. US20020082216A1
; GENERAL INFORMATION:
; APPLICANT: Fanger, Gary R.
; APPLICANT: Foy, Theresia M.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Reed, Steven G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; FILE REFERENCE: 210121.479C1
; CURRENT APPLICATION NUMBER: US/09/757,417
; CURRENT FILING DATE: 2001-01-08
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-757-417-46

Alignment Scores:
Pred. No.:      2,44e-16      Length:      399
Score:          148.00      Matches:      30
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:      0
DB:              9          Gaps:      0

US-09-684-215A-17 (1-30) x US-09-757-417-46 (1-399)
QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20
Db 82 GGCAGGCGATGGCGATCGCGGCCAGATC 111
```



```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/084,843
FILING DATE: 25-Feb-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/072,967
FILING DATE: 05-MAY-1998
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.411C9
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 447 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-084-843-4

Alignment Scores:
Pred. No.:      2.78e-16      Length:      447
Score:          148.00       Matches:     30
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches:    0
Query Match: .        100.00% Indels:         0
DB:                12       Gaps:         0

US-09-684-215A-17 (1-30) x US-10-084-843-4 (1-447)

QY      1 ThrAlaIaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPhe
Db      11 ACGGCGGGTTCGATACCTCCAGCTGTCCAGGGTGCGCAGGATG
QY      21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
Db      71 GGCGAGCGCATGGCGATCGCGGCCAGATC 100

RESULT 7
US-10-193-002-4
Sequence 4, Application US/10193002
Publication No. US20030135026A1
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
Skelky, Yasir A.W.
Dillon, Davin C.
Campos-Neto, Antonia
Houghton, Raymond
Vedvick, Thomas S.
Twardzik, Daniel R.
Lodes, Michael J.
Hendrickson, Ronald C.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSTIC
TUBERCULOSIS
NUMBER OF SEQUENCES: 350
CORRESPONDENCE ADDRESS:
ADDRESS: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

```

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/193,002  
FILING DATE: 10-Jul-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/072,596  
FILING DATE: 05-MAY-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Makí, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.417C9  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 447 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-10-193-002-4

Alignment Scores:  
Pred. No.: 2,78e-16 Length: 447  
Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 12 Gaps:

US-09-684-215A-17 (1-30) x US-10-193-002-4 (1-447)

QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20  
|||||  
Db II ACGGCGCGGTCGGAATACTTCCAGCTGTCCTCCAGGTTGGCAGGAGTCCGCAATCCGATC 70  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30  
|||||  
Db 71 GGCAGGCGGATGGCGATCGCGGCCAGATC 100

## RESULT 8

US-09-759-143-822  
; Sequence 822, Application US/09759143  
; Patent No. US200202248A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun C.  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuqui  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedwick, Thomas S.  
; APPLICANT: Carter, Darrick  
; APPLICANT: Li, Samuel  
; APPLICANT: Wang, Aljun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Hepler, William  
; APPLICANT: Hural, John  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Houghton, Raymond L.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.427C24  
; CURRENT FILING DATE: 2001-01-12  
; CURRENT APPLICATION NUMBER: US/09/759,143  
; NUMBER OF SEQ ID NOS: 934  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 822  
; LENGTH: 675  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-780-669-822

Alignment Scores:  
Pred. No.: 4.47e-16 Length: 675  
Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps:

US-09-684-215A-17 (1-30) x US-09-780-669-822 (1-675)

QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20  
|||||

; LENGTH: 675  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-759-143-822  
Alignment Scores:  
Pred. No.: 4.47e-16 Length: 675  
Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps:

US-09-684-215A-17 (1-30) x US-09-759-143-822 (1-675)

QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20  
|||||  
Db 22 ACGGCGCGGTCGGAATACTTCCAGCTGTCCTCCAGGTTGGCAGGAGTCCGCAATCCGATC 81  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30  
|||||  
Db 82 GGCAGGCGGATGGCGATCGCGGCCAGATC 111

## RESULT 9

US-09-780-669-822  
; Sequence 822, Application US/09780669  
; Patent No. US20020051977A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun C.  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuqui  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedwick, Thomas S.  
; APPLICANT: Carter, Darrick  
; APPLICANT: Li, Samuel  
; APPLICANT: Wang, Aljun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Hepler, William  
; APPLICANT: Hural, John  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Houghton, Raymond L.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.427C24  
; CURRENT FILING DATE: 2001-02-09  
; CURRENT APPLICATION NUMBER: US/09/780,669  
; NUMBER OF SEQ ID NOS: 943  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 822  
; LENGTH: 675  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-780-669-822

Alignment Scores:  
Pred. No.: 4.47e-16 Length: 675  
Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps:

US-09-684-215A-17 (1-30) x US-09-780-669-822 (1-675)

QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20  
|||||

```

; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.534C2

```



```

; FILE REFERENCE: 014058-009020US
; CURRENT APPLICATION NUMBER: US/09/287,849
; CURRENT FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: US 08/818,112
; PRIOR FILING DATE: 1997-03-13
; PRIOR APPLICATION NUMBER: US 08/942,578
; PRIOR FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: US 09/025,197
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 09/056,556
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 702
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:bi-fusion
; OTHER INFORMATION: protein Ral2-DPPD (designated Mtb24), reading
; OTHER INFORMATION: frame 1
; NAME/KEY: CDS
; LOCATION: (1)..(693)
; OTHER INFORMATION: bi-fusion protein Ral2-DPPD (designated Mtb24),
; OTHER INFORMATION: reading frame 1
; NAME/KEY: CDS
; LOCATION: (2)..(700)
; OTHER INFORMATION: reading frame 2
; NAME/KEY: CDS
; LOCATION: (3)..(701)
; OTHER INFORMATION: reading frame 3
US-09-287-849-27

Alignment Scores:
Pred. No.: 4.68e-16 Length: 702
Score: 148.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-684-215A-17 (1-30) x US-09-287-849-27 (1-702)

Qy 1 ThrAlaAlaSerAspPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 25 ACGGCGGCTCGGATAACTTCCAGCTGCCAGGGTGGCAGGGATTCCGCATTCGCATC 84
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
Db 85 GGGCAGCGGATGGCGATCGGGGCCAGATC 114

Search completed: September 5, 2003, 12:04:44
Job time : 44.8851 secs

```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 5, 2003, 07:52:39 ; Search time 535.655 seconds  
(without alignments)  
2291.192 Million cell updates/sec

Title: US-09-684-215A-17  
Perfect score: 148  
Sequence: 1 TAASDNFQLSQGGQGAIPIGQAMAIAGQI 30

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlp  
-O=/cgn2\_1/USPTO.spool\_p/US09684215/runat\_05092003\_072202\_559/app\_query.fasta\_1.853  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=Dits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pcpt -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09684215 @CGN\_1\_1\_3962 @runat\_05092003\_072202\_559 -NCPU=3  
-NO\_MMAB -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_ov.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_om.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_ph.\*  
24: em\_pl.\*  
25: em\_ro.\*  
26: em\_ro.\*  
27: em\_ro.\*  
28: em\_un.\*

29: em\_vi.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pln.\*  
35: em\_htg\_rdd.\*  
36: em\_htg\_mam.\*  
37: em\_htg\_vrt.\*  
38: em\_sy.\*  
39: em\_htgo\_hum.\*  
40: em\_htgo\_mus.\*  
41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	148	100.0	186	6	AX369169 Sequence
2	148	100.0	447	6	ARI69152 Sequence
3	148	100.0	447	6	ARI82442 Sequence
4	148	100.0	447	6	ARI194825 Sequence
5	148	100.0	447	6	AR233097 Sequence
6	148	100.0	447	6	AX429596 Sequence
7	148	100.0	447	6	BD006325 Compounds
8	148	100.0	447	6	BD006445 Compounds
9	148	100.0	447	6	BD069285 Compounds
10	148	100.0	675	6	AR261272 Sequence
11	148	100.0	675	6	AX201049 Sequence
12	148	100.0	675	6	AX267848 Sequence
13	148	100.0	822	6	AR277645 Sequence
14	148	100.0	822	6	AX369152 Sequence
15	148	100.0	861	6	AX369167 Sequence
16	148	100.0	894	6	AX351489 Sequence
17	148	100.0	900	6	AR220690 Sequence
18	148	100.0	900	6	AX365960 Sequence
19	148	100.0	915	6	AR261281 Sequence
20	148	100.0	915	6	AX201061 Sequence
21	148	100.0	915	6	AX267860 Sequence
22	148	100.0	945	6	AR277644 Sequence
23	148	100.0	945	6	AX369151 Sequence
24	148	100.0	1012	6	AR220689 Sequence
25	148	100.0	1012	6	AX365958 Sequence
26	148	100.0	1068	6	AX005788 Sequence
27	148	100.0	1143	6	AX005790 Sequence
28	148	100.0	1155	6	AX369165 Sequence
29	148	100.0	1203	6	AX201078 Sequence
30	148	100.0	1203	6	AX267877 Sequence
31	148	100.0	1464	6	AR229410 Sequence
32	148	100.0	1464	6	AX156105 Sequence
33	148	100.0	1464	6	AX361955 Sequence
34	148	100.0	1557	6	AR229398 Sequence
35	148	100.0	1557	6	AX156089 Sequence
36	148	100.0	1557	6	AX361939 Sequence
37	148	100.0	1578	6	AR229328 Sequence
38	148	100.0	1578	6	AX155945 Sequence
39	148	100.0	1578	6	AX361795 Sequence
40	148	100.0	1590	6	AX316986 Sequence
41	148	100.0	1752	6	AR229413 Sequence
42	148	100.0	1752	6	AX156109 Sequence
43	148	100.0	1752	6	AX361959 Sequence
44	148	100.0	1758	6	AR229401 Sequence
45	148	100.0	1758	6	AX156093 Sequence

ALIGNMENTS

AX369169 186 bp DNA linear PAT 16-FEB-2002  
LOCUS Sequence 1879 from Patent WO0204514.  
ACCESSION AX369169  
VERSION AX369169.1 GI:18857178  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
REFERENCE Wang, T., Watanabe, Y., Henderson, R.A., Johnson, J.C., Retter, M.W.,  
Mannerakis, M., Carter, D., Fanger, G.R., Vedvick, T.S., Bangur, C.S.,  
McNabb, A., Fanger, N., Switzer, A., McNeill, P.D. and Clapper, J.D.  
TITLE Compositions and methods for the therapy and diagnosis of lung  
cancer  
JOURNAL Patent: WO 0204514-A 1879 17-JAN-2002;  
CORIXA CORPORATION (US)  
FEATURES  
source  
1. 186  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 39 a 61 c 47 g 39 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 3,84e-14 Length: 186  
Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-09-684-215A-17 (1-30) x AX369169 (1-186)  
QY 1 ThrAlaAlaSerAspAsnGlnLeuSerGlnGlyGlnGlyGlnGlyPheAlaIleProIle 20  
Db 22 ACGGCCGGCTGCGATAACTTCCAGCTGTCCAGGGTGGCAGGGATTCCGCCATTCGGATC 81  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30  
Db 82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111  
RESULT 2  
AR169152  
LOCUS AR169152  
DEFINITION Sequence 4 from patent US 6290969.  
ACCESSION AR169152  
VERSION AR169152.1 GI:17906927  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
AUTHORS 1 (bases 1 to 447)  
TITLE Reed, S.G., Skeiky, Y.A.W., Dillon, D.C., Campos-Neto, A., Houghton, R.,  
Vedvick, T.S. and Twardzik, D.R.  
JOURNAL Compositions and methods for immunotherapy and diagnosis of  
tuberculosis  
FEATURES Patent: US 6290969-A 4 18-SEP-2001;  
source Location/Qualifiers  
1. 447  
/organism="unknown"  
BASE COUNT 79 a 146 c 149 g 72 t 1 others  
ORIGIN  
Alignment Scores:  
Pred. No.: 9.83e-14 Length: 447  
Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-09-684-215A-17 (1-30) x AR169152 (1-447)  
QY 1 ThrAlaAlaSerAspAsnGlnLeuSerGlnGlyGlnGlyGlnGlyPheAlaIleProIle 20  
Db 11 ACGGCCGGCTGCGATAACTTCCAGCTGTCCAGGGTGGCAGGGATTCCGCCATTCGGATC 70  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30  
Db 71 GGGCAGGCGATGGCGATCGCGGGCCAGATC 100  
RESULT 4  
AR194825  
LOCUS AR194825  
DEFINITION Sequence 4 from patent US 6350456.  
ACCESSION AR194825  
VERSION AR194825.1 GI:20244262  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
AUTHORS 1 (bases 1 to 447)  
TITLE Reed, S.G., Skeiky, Y.A.W. and Dillon, D.C.  
JOURNAL Compositions and methods for the prevention and treatment of M.  
tuberculosis infection  
FEATURES Patent: US 6350456-A 4 26-FEB-2002;  
source Location/Qualifiers  
1. 447  
/organism="unknown"  
BASE COUNT 79 a 146 c 149 g 72 t 1 others  
ORIGIN  
Alignment Scores:  
Pred. No.: 9.83e-14 Length: 447



Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-684-215A-17 (1-30) x AR194825 (1-447)

Qy 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyGlnGlyPheAlaIleProIle 20  
|||||  
Db 11 ACGGCCGCGTCCGATACCTTCCAGCTGTCCAGGGTGGCAGGATTCCGCATTCGCATC 70  
Qy 21 GlyClnAlaMetAlaIleAlaGlyGlnIle 30  
|||||  
Db 71 GGCAGCGGATGGCGATCGCGGGCCAGATC 100

#### RESULT 5

AR233097 AR233097 447 bp DNA linear PAT 20-DEC-2002  
LOCUS  
DEFINITION Sequence 4 from patent US 6458366.  
ACCESSION AR233097  
VERSION AR233097.1 GI:27275533  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 447)  
AUTHORS Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Campos-Neto,A., Houghton,R.,  
Vedvick,T.S., Twardzik,D.R., Lodes,M.J. and Hendrickson,R.C.

TITLE Compounds and methods for diagnosis of tuberculosis  
JOURNAL Patent: US 6458366-A 4 01-OCT-2002;  
FEATURES Location/Qualifiers  
source 1. .447  
BASE COUNT 79 a 146 c 149 g 72 t 1 others  
ORIGIN

Alignment Scores:  
Pred. No.: 9.83e-14 Length: 447  
Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-684-215A-17 (1-30) x AR233097 (1-447)

Qy 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyGlnGlyPheAlaIleProIle 20  
|||||  
Db 11 ACGGCCGCGTCCGATACCTTCCAGCTGTCCAGGGTGGCAGGATTCCGCATTCGCATC 70  
Qy 21 GlyClnAlaMetAlaIleAlaGlyGlnIle 30  
|||||  
Db 71 GGCAGCGGATGGCGATCGCGGGCCAGATC 100

#### RESULT 6

AX429596 AX429596 447 bp DNA linear PAT 21-JUN-2002  
LOCUS  
DEFINITION Sequence 4 from Patent EP1203817.  
ACCESSION AX429596  
VERSION AX429596.1 GI:21540845  
KEYWORDS  
SOURCE unidentified  
ORGANISM unclassified.

REFERENCE 1  
AUTHORS Reed,S.G., Skeiky,Y.A., Dillon,D.C., Campos-Neto,A., Houghton,R.L.,  
Vedvick,T.S. and Twardzik,D.R.

TITLE Compounds and methods for immunotherapy and diagnosis of  
JOURNAL tuberculosis  
Patent: EP 1203817-A 4 08-MAY-2002;  
CORIXA CORPORATION (US)  
FEATURES Location/Qualifiers

source 1. .447  
BASE COUNT 79 a 146 c 149 g 72 t 1 others  
ORIGIN

Alignment Scores:  
Pred. No.: 9.83e-14 Length: 447  
Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-684-215A-17 (1-30) x AX429596 (1-447)

Qy 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyGlnGlyPheAlaIleProIle 20  
|||||  
Db 11 ACGGCCGCGTCCGATACCTTCCAGCTGTCCAGGGTGGCAGGATTCCGCATTCGCATC 70  
Qy 21 GlyClnAlaMetAlaIleAlaGlyGlnIle 30  
|||||  
Db 71 GGCAGCGGATGGCGATCGCGGGCCAGATC 100

#### RESULT 7

BD006325 BD006325 447 bp DNA linear PAT 31-JAN-2002  
LOCUS  
DEFINITION Compounds and methods for diagnosis of Tuberculosis.

ACCESSION BD006325  
VERSION BD006325.1 GI:18634696  
KEYWORDS JP 2001500383-A/4.  
SOURCE unidentified

ORGANISM unclassified  
REFERENCE 1 (bases 1 to 447)

AUTHORS Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Neto,A.C., Houghton,R.,  
Vedvick,T.S., Twardzik,D.R. and Lodes,M.J.  
TITLE Compounds and methods for diagnosis of Tuberculosis  
JOURNAL Patent: JP 2001500383-A 4 16-JAN-2001;  
CORIXA CORP

COMMENT OS Unidentified  
PN JP 2001500383-A/4  
PD 16-JAN-2001  
PF 07-OCT-1997 JP 1998518432  
PR 11-OCT-1996 US 08/729622,13-MAR-1997 US 08/818111 PI

STEVEN G REED, YASIR A W SKEIKY, DAVIN C DILLON, PI ANTONIO CAMPOS  
NETO,  
PI RAYMOND HOUGHTON, THOMAS S VEDVICK, DANIEL R TWARDZIK, PI  
MICHAEL J LODES  
PC C12N15/31.C07K14/35.C07K16/12.C12Q1/68.C12N15/62.G01N33/53 CC  
Strandedness: Single;  
CC Topology: Linear;  
FH Key Location/Qualifiers  
FT source 1. .447  
FT /organism='Unidentified'.

FEATURES  
source 1. .447  
Location/Qualifiers  
/organism='unidentified'  
/mol\_type='genomic DNA'  
/db\_xref='taxon:32644'

BASE COUNT 79 a 146 c 149 g 72 t 1 others  
ORIGIN

Alignment Scores:  
Pred. No.: 9.83e-14 Length: 447  
Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-684-215A-17 (1-30) x BD006325 (1-447)

```

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyGlnGlyPheAlaIleProfile 20
Db 11 ACGGCGCGTCCGATAACTCCAGCTGTCAGGGTGGCAGGATTCGCCATTCGCATC 70
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
Db 71 GGGCAGCGGATGGCGATCGCGGCCAGATC 100

RESULT 8
BD006445
LOCUS
DEFINITION
  Compounds and methods for immunotherapy and diagnosis of
  Tuberculosis.
ACCESSION
  BD006445
VERSION
  JP 2001501832-A/4.
KEYWORDS
  unidentified
SOURCE
  unclassified
ORGANISM
  unclassified
REFERENCE
  1 (bases 1 to 447)
AUTHORS
  Reed, S.G., Skeiky, Y.A.W., Dillon, D.C., Neto, A.C., Houghton, R.,
  Vedvick, T.S., Twardzik, D.R. and Lodes, M.J.
TITLE
  Compounds and methods for immunotherapy and diagnosis of
  Tuberculosis.
JOURNAL
  CORIXA CORP
COMMENT
  OS Unidentified
  PN JP 2001501832-A/4
  PD 13-FEB-2001
  PF 07-OCT-1997 JP 1998518456
  PR 11-OCT-1996 US 08/730510, 13-MAR-1997 US 08/818112 PI
  STEVEN G REED, YASIR A W SKEIKY, DAVIN C DILLON, PI ANTONIO CAMPOS
  NETO,
  PI RAYMOND HOUGHTON, THOMAS S VEDVICK, DANIEL R TWARDZIK, PI
  MICHAEL J LODES
  PC C12N15/31, C07K14/35, A61K39/04, A61K48/00, A61K49/00, C12N15/62,
  PC C07K19/00,
  PC G01N33/50, G01N33/60, G01N33/569, C12N1/19, C12N1/20, C12N1/21, PC
  C12N5/10//
  PC (C12N1/21, C12R1:19)
  CC Strandedness: Single;
  CC Topology: Linear;
  FH Key Location/Qualifiers
  FT source
  FT Location/Qualifiers
  1. .447
  /organism="Unidentified"
  /organism="unidentified"
  /mol_type="genomic DNA"
  /db_xref="taxon:32644"
  /db_xref="taxon:32644"

BASE COUNT 79 a 146 c 149 g 72 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 9,83e-14 Length: 447
Score: 148.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-684-215A-17 (1-30) x BD006445 (1-447)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyGlnGlyPheAlaIleProfile 20
Db 11 ACGGCGCGTCCGATAACTCCAGCTGTCAGGGTGGCAGGATTCGCCATTCGCATC 70
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
Db 71 GGGCAGCGGATGGCGATCGCGGCCAGATC 100

RESULT 9
BD0069285
LOCUS
DEFINITION
  Compounds and methods for immunotherapy and diagnosis of
  Tuberculosis.
ACCESSION
  BD0069285
VERSION
  JP 2001517069-A/4.
KEYWORDS
  unidentified
SOURCE
  unclassified
ORGANISM
  unclassified
REFERENCE
  1 (bases 1 to 447)
AUTHORS
  Reed, S.G., Skeiky, Y.A.W., Dillon, D.C., Neto, A.C., Houghton, R.,
  Vedvick, T.H. and Twardzik, D.R.
TITLE
  Compounds and methods for immunotherapy and diagnosis of
  Tuberculosis.
JOURNAL
  CORIXA CORP
COMMENT
  OS Unidentified
  PN JP 2001517069-A/4
  PD 02-OCT-2001
  PF 30-AUG-1996 JP 1997511464
  PR 01-SEP-1995 US 08/523436, 22-SEP-1995 US 08/533634 PR
  22-MAR-1996 US 08/620874, 05-JUN-1996 US 08/659683 PR
  12-JUL-1996 US 08/680574
  PI STEVEN G REED, YASIR A W SKEIKY, DAVIN C DILLON, ANTONIO CAMPOS
  PI NETO,
  PI RAYMOND HOUGHTON, THOMAS H VEDVICK, DANIEL R TWARDZIK PC
  C12N15/31, C07K14/35, A61K38/16, C12N15/62, G01N33/569, C12Q1/68, PC
  C12N5/10,
  PC C12N1/21//A61K39/04, C12N1/21, C12R1:19)
  CC Strandedness: Single;
  CC Topology: Linear;
  CC Compounds and methods for immunotherapy and diagnosis of CC
  Tuberculosis
  FH Key Location/Qualifiers
  FT source
  FT Location/Qualifiers
  1. .447
  /organism="Unidentified"
  /organism="unidentified"
  /mol_type="genomic DNA"
  /db_xref="taxon:32644"
  /db_xref="taxon:32644"

BASE COUNT 79 a 146 c 149 g 72 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 9,83e-14 Length: 447
Score: 148.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-684-215A-17 (1-30) x BD069285 (1-447)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyGlnGlyPheAlaIleProfile 20
Db 11 ACGGCGCGTCCGATAACTCCAGCTGTCAGGGTGGCAGGATTCGCCATTCGCATC 70
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
Db 71 GGGCAGCGGATGGCGATCGCGGCCAGATC 100

RESULT 10
AR261272
LOCUS
DEFINITION
  Sequence 822 from patent US 6321716.
ACCESSION
  AR261272
VERSION
  AR261272.1 GI:28072035
KEYWORDS
  Unknown.
SOURCE
  Unclassified.
REFERENCE
  1 (bases 1 to 675)
AUTHORS
  Mashiki, Z., and Harada, J.

```

```

LOCUS
DEFINITION
  Compounds and methods for immunotherapy and diagnosis of
  Tuberculosis.
ACCESSION
  BD069285
VERSION
  JP 2001517069-A/4.
KEYWORDS
  unidentified
SOURCE
  unclassified
ORGANISM
  unclassified
REFERENCE
  1 (bases 1 to 447)
AUTHORS
  Reed, S.G., Skeiky, Y.A.W., Dillon, D.C., Neto, A.C., Houghton, R.,
  Vedvick, T.H. and Twardzik, D.R.
TITLE
  Compounds and methods for immunotherapy and diagnosis of
  Tuberculosis.
JOURNAL
  CORIXA CORP
COMMENT
  OS Unidentified
  PN JP 2001517069-A/4
  PD 02-OCT-2001
  PF 30-AUG-1996 JP 1997511464
  PR 01-SEP-1995 US 08/523436, 22-SEP-1995 US 08/533634 PR
  22-MAR-1996 US 08/620874, 05-JUN-1996 US 08/659683 PR
  12-JUL-1996 US 08/680574
  PI STEVEN G REED, YASIR A W SKEIKY, DAVIN C DILLON, ANTONIO CAMPOS
  PI NETO,
  PI RAYMOND HOUGHTON, THOMAS H VEDVICK, DANIEL R TWARDZIK PC
  C12N15/31, C07K14/35, A61K38/16, C12N15/62, G01N33/569, C12Q1/68, PC
  C12N5/10,
  PC C12N1/21//A61K39/04, C12N1/21, C12R1:19)
  CC Strandedness: Single;
  CC Topology: Linear;
  CC Compounds and methods for immunotherapy and diagnosis of CC
  Tuberculosis
  FH Key Location/Qualifiers
  FT source
  FT Location/Qualifiers
  1. .447
  /organism="Unidentified"
  /organism="unidentified"
  /mol_type="genomic DNA"
  /db_xref="taxon:32644"
  /db_xref="taxon:32644"

BASE COUNT 79 a 146 c 149 g 72 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 9,83e-14 Length: 447
Score: 148.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-684-215A-17 (1-30) x BD069285 (1-447)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyGlnGlyPheAlaIleProfile 20
Db 11 ACGGCGCGTCCGATAACTCCAGCTGTCAGGGTGGCAGGATTCGCCATTCGCATC 70
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
Db 71 GGGCAGCGGATGGCGATCGCGGCCAGATC 100

RESULT 10
AR261272
LOCUS
DEFINITION
  Sequence 822 from patent US 6321716.
ACCESSION
  AR261272
VERSION
  AR261272.1 GI:28072035
KEYWORDS
  Unknown.
SOURCE
  Unclassified.
REFERENCE
  1 (bases 1 to 675)
AUTHORS
  Mashiki, Z., and Harada, J.

```

TITLE Negative pressure control apparatus for engine mounted in vehicle  
JOURNAL Patent: US 6321716-A 822 27-NOV-2001;

FEATURES  
source Location/Qualifiers  
1..675

BASE COUNT 162 a 197 c 190 g 126 t  
ORIGIN

Alignment Scores:  
Pred. No.: 1.53e-13 Length: 675  
Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-684-215A-17 (1-30) x AR261272 (1-675)

Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
|||||  
Db 22 ACGGCGCGTCCGATACCTCCAGCTGTCCAGGGTGGCAGGATTCCGCATTCCGATC 81  
|||||  
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30  
|||||  
Db 82 GGCAGGCGATGCGATCGCGGCCAGATC 111  
|||||

RESULT 11

AX201049  
LOCUS AX201049 675 bp DNA linear PAT 29-AUG-2001

DEFINITION Sequence 679 from Patent WO0151633.

ACCESSION AX201049

VERSION AX201049.1 GI:15390857

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS

Xu, J., Dillon, D.C., Mitcham, J.L., Harlocker, S.L., Jiang, Y.,  
Reed, S.G., Kalos, M.D., Fanger, G.R., Day, C.H., Retter, M.W.,  
Stolk, J.A., Skeiky, Y.A., Wang, A. and Weagher, M.J.  
Compositions and methods for the therapy and diagnosis of prostate  
cancer

JOURNAL Patent: WO 0151633-A 679 19-JUL-2001;

CORIXA CORPORATION (US)

FEATURES  
source Location/Qualifiers

1..675

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

BASE COUNT 162 a 197 c 190 g 126 t

ORIGIN

Alignment Scores:  
Pred. No.: 1.53e-13 Length: 675  
Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-684-215A-17 (1-30) x AX201049 (1-675)

Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
|||||  
Db 22 ACGGCGCGTCCGATACCTCCAGCTGTCCAGGGTGGCAGGATTCCGCATTCCGATC 81  
|||||

Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30  
|||||

Db 82 GGCAGGCGATGCGATCGCGGCCAGATC 111  
|||||

RESULT 12

AX267848

LOCUS AX267848 822 bp DNA linear PAT 26-OCT-2001

DEFINITION Sequence 822 from Patent WO0173032.

ACCESSION AX267848

VERSION AX267848.1 GI:16516494

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS

Xu, J., Dillon, D.C., Mitcham, J.L., Harlocker, S.L., Jiang, Y.,  
Kalos, M.D., Fanger, G.R., Retter, M.W., Stolk, J.A., Day, C.H.,  
Vedvick, T.S., Carter, D., Li, S.X., Wang, A., Skeiky, Y.A., Hepler, W.T.  
and Henderson, R.A.  
Compositions and methods for the therapy and diagnosis of prostate  
cancer

JOURNAL Patent: WO 0173032-A 822 04-OCT-2001;

CORIXA CORPORATION (US)

FEATURES  
source Location/Qualifiers

1..675

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

BASE COUNT 162 a 197 c 190 g 126 t

ORIGIN

Alignment Scores:

Pred. No.: 1.53e-13 Length: 675  
Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-684-215A-17 (1-30) x AX267848 (1-675)

Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
|||||  
Db 22 ACGGCGCGTCCGATACCTCCAGCTGTCCAGGGTGGCAGGATTCCGCATTCCGATC 81  
|||||

Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30  
|||||

Db 82 GGCAGGCGATGCGATCGCGGCCAGATC 111  
|||||

RESULT 13

AX277645

LOCUS AR277645 822 bp DNA linear PAT 10-APR-2003

DEFINITION Sequence 1862 from patent US 6509448.

ACCESSION AR277645

VERSION AR277645.1 GI:29711294

KEYWORDS

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 822)

AUTHORS

Wang, T., Bangur, C.S., Lodes, M.J., Fanger, G.R., Vedvick, T.S.,  
Carter, D., Retter, M.W., Mannion, J., Fan, L. and Wang, A.  
Compositions and methods for the therapy and diagnosis of lung  
cancer

JOURNAL Patent: US 6509448-A 1862 21-JAN-2003;

FEATURES  
source Location/Qualifiers

1..822

/organism="unknown"

BASE COUNT 146 a 281 c 270 g 125 t

ORIGIN

Alignment Scores:

Pred. No.: 1.89e-13 Length: 822  
Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-684-215A-17 (1-30) x AR277645 (1-822)

Qy 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20  
|||||  
Db 22 ACGGCGCGGTCCGATAACTTCCAGCTGTCCAGGGTGGCAGGGATTGCCATTCGGATC 81  
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30  
|||||  
Db 82 GGGCAGCGGATGGCGATCGCGGCCAGATC 111

## RESULT 14

AX369152

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

BASE COUNT

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-09-684-215A-17 (1-30) x AX369152 (1-822)

Qy 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20

Db 22 ACGGCGCGGTCCGATAACTTCCAGCTGTCCAGGGTGGCAGGGATTGCCATTCGGATC 81

Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30

Db 82 GGGCAGCGGATGGCGATCGCGGCCAGATC 111

RESULT 15

AX369167

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

BASE COUNT

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-09-684-215A-17 (1-30) x AX369167 (1-822)

Qy 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20

Db 22 ACGGCGCGGTCCGATAACTTCCAGCTGTCCAGGGTGGCAGGGATTGCCATTCGGATC 81

Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30

Db 82 GGGCAGCGGATGGCGATCGCGGCCAGATC 111

RESULT 16

AX369167

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

BASE COUNT

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-09-684-215A-17 (1-30) x AX369167 (1-822)

Qy 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20

Db 22 ACGGCGCGGTCCGATAACTTCCAGCTGTCCAGGGTGGCAGGGATTGCCATTCGGATC 81

Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30

Db 82 GGGCAGCGGATGGCGATCGCGGCCAGATC 111

Search completed: September 5, 2003, 10:31:40

Job time : 536.989 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 5, 2003, 08:28:59 ; Search time 16.0345 Seconds  
(without alignments)  
825.814 Million cell updates/sec

Title: us-09-684-215a-17

Perfect score: 148

Sequence: 1 TAASDNFLSQSGGFAIPICQAMAIAGQI 30

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-O=/cgn2\_1/USPTO\_spool\_p/US09684215/runat\_05092003\_072203\_593/app-query.fasta\_1.853  
-DB-Issued\_Patents\_NA -QPMF-fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=BIOSUM62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09684215 -CGN\_1\_145\_runat\_05092003\_072203\_593 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued\_Patents\_NA: \*  
1: /cgn2\_1/ptodata/1/ina/5A\_COMB.seq: \*  
2: /cgn2\_1/ptodata/1/ina/5B\_COMB.seq: \*  
3: /cgn2\_1/ptodata/1/ina/6A\_COMB.seq: \*  
4: /cgn2\_1/ptodata/1/ina/6B\_COMB.seq: \*  
5: /cgn2\_1/ptodata/1/ina/PCTUS\_COMB.seq: \*  
6: /cgn2\_1/ptodata/1/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	148	100.0	447	3	US-08-818-111-4
2	148	100.0	447	4	US-08-818-111-4
3	148	100.0	447	4	US-09-056-556-4
4	148	100.0	447	4	US-09-072-596-4
5	148	100.0	822	4	US-09-736-457-1862
6	148	100.0	900	4	US-09-643-597-353
7	148	100.0	900	4	US-09-606-421B-353
8	148	100.0	945	4	US-09-736-457-1861
9	148	100.0	1012	4	US-09-643-597-351
10	148	100.0	1012	4	US-09-606-421B-351
11	148	100.0	1464	4	US-09-620-412C-348
12	148	100.0	1464	4	US-09-598-419-348

Sequence 332, App  
Sequence 332, App  
Sequence 188, App  
Sequence 188, App  
Sequence 188, App  
Sequence 188, App  
Sequence 332, App  
Sequence 332, App  
Sequence 336, App  
Sequence 308, App  
Sequence 308, App  
Sequence 17, Appl  
Sequence 17, Appl  
Sequence 17, Appl  
Sequence 324, App  
Sequence 324, App  
Sequence 316, App  
Sequence 316, App  
Sequence 340, App  
Sequence 340, App  
Sequence 356, App  
Sequence 356, App  
Sequence 312, App  
Sequence 312, App  
Sequence 344, App  
Sequence 344, App  
Sequence 320, App  
Sequence 320, App  
Sequence 328, App  
Sequence 320, App  
Sequence 328, App  
Sequence 1, Appl  
Sequence 2, Appl

#### ALIGNMENTS

RESULT 1  
US-08-818-112-4  
; Sequence 4, Application US/08818112  
; Patent No. 6290959  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Campos-Neto, Antonio  
; APPLICANT: Houghton, Raymond  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Twardzik, Daniel R.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS  
; NUMBER OF SEQUENCES: 153  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED AND BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; City: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; .COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/818,112  
; FILING DATE: 13-MAR-1997  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Makl, David J.  
; REGISTRATION NUMBER: 31,392  
; REFERENCE/DOCKET NUMBER: 210121.411C6  
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 447 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-818-112-4

## Alignment Scores:

Pred. No.: 9.5e-16 Length: 447  
Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0

US-09-684-215A-17 (1-30) x US-08-818-112-4 (1-447)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20  
|||||  
Db 11 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGTGGCAGGATTCGCCATTCCGATC 70  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30  
|||||  
Db 71 GGGCAGGCGATGCGATCGCGGCCAGATC 100

## RESULT 2

US-08-818-111-4

Sequence 4, Application US/08818111

Patent No. 6338852

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.

APPLICANT: Skeiky, Yasir A.W.

APPLICANT: Dillon, Davin C.

APPLICANT: Campos-Neto, Antonia

APPLICANT: Houghton, Raymond

APPLICANT: Vedvick, Thomas S.

APPLICANT: Twardzik, Daniel R.

TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF TUBERCULOSIS

NUMBER OF SEQUENCES: 148

CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED and BERRY LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/818,111

FILING DATE: 13-MAR-1997

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Maki, David J.

REGISTRATION NUMBER: 31,392

REFERENCE/DOCKET NUMBER: 210121.417C6

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 447 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-818-111-4

## Alignment Scores:

Pred. No.: 9.5e-16 Length: 447  
Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-684-215A-17 (1-30) x US-08-818-111-4 (1-447)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20  
|||||  
Db 11 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGTGGCAGGATTCGCCATTCCGATC 70  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30  
|||||  
Db 71 GGGCAGGCGATGCGATCGCGGCCAGATC 100

## RESULT 3

US-09-056-556-4

Sequence 4, Application US/09056556

Patent No. 6350456

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.

APPLICANT: Skeiky, Yasir A.W.

APPLICANT: Dillon, Davin C.

TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND

NUMBER OF SEQUENCES: 241

CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED and BERRY LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/056,556

FILING DATE: 07-APR-1998

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Maki, David J.

REGISTRATION NUMBER: 31,392

REFERENCE/DOCKET NUMBER: 210121.457

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 447 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-09-056-556-4

## Alignment Scores:

Pred. No.: 9.5e-16 Length: 447  
Score: 148.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-684-215A-17 (1-30) x US-09-056-556-4 (1-447)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20  
|||||  
Db 11 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGTGGCAGGATTCGCCATTCCGATC 70



[illegible]

US-09-684-215A-17 (1-30) x US-09-643-597-353 (1-900)					
Score:	148.00	Matches:	30		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	100.00%	Indels:	0		
DB:	4	Gaps:	0		
US-09-684-215A-17 (1-30) x US-09-643-597-353 (1-900)					
Qy	1	ThrAlaAlaSerAspAsnPhgInLeuSerGlnGlycInGlyPheAlaIleProfile	20		
Dd	22	ACGGCCCGTCCGATAACTTCAGCTGTCCAGGTTGGCAGGATTCGCCATTCCGATC	81		
Qy	21	GlyGlnAlaMetAlaIleAlaGlyGlnIle	30		
Dd	82	GGCAGCGCATGCAGTCCGGGCCAGATC	111		
RESULT 7					
US-09-606-421B-353					
; Sequence 353, Application US/09606421B					
; Patent No. 6531315					
; GENERAL INFORMATION:					
; APPLICANT: Wang, Tongtong					
; APPLICANT: Fan, Liqun					
; APPLICANT: Kalos, Michael D.					
; APPLICANT: Bangur, Chaitanya S.					
; APPLICANT: Hosken, Nancy					
; APPLICANT: Fanger, Gary R.					
; APPLICANT: Li, Samuel X.					
; APPLICANT: Wang, Aijun					
; APPLICANT: Skeiky, Yasir A.W.					
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY					
; FILE REFERENCE: 210121.455C9					
; CURRENT APPLICATION NUMBER: US/09/606,421B					
; CURRENT FILING DATE: 2000-06-28					
; NUMBER OF SEQ ID NOS: 358					
; SOFTWARE: FastSeq for Windows Version 3.0					
; SEQ ID NO 353					
; LENGTH: 900					
; TYPE: DNA					
; ORGANISM: Homo sapiens					
US-09-606-421B-353					
Alignment Scores:					
Pred. No.:	2,27e-15	Length:	900		
Score:	148.00	Matches:	30		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	100.00%	Indels:	0		
DB:	4	Gaps:	0		
US-09-684-215A-17 (1-30) x US-09-606-421B-353 (1-900)					
Qy	1	ThrAlaAlaSerAspAsnPhgInLeuSerGlnGlycInGlyPheAlaIleProfile	20		
Dd	22	ACGGCCCGTCCGATAACTTCAGCTGTCCAGGTTGGCAGGATTCGCCATTCCGATC	81		
Qy	21	GlyGlnAlaMetAlaIleAlaGlyGlnIle	30		
Dd	82	GGCAGCGCATGCAGTCCGGGCCAGATC	111		
RESULT 8					
US-09-736-457-1861					
; Sequence 1861, Application US/09736457					
; Patent No. 6509448					
; GENERAL INFORMATION:					
; APPLICANT: Wang, Tongtong					
; APPLICANT: Bangur, Chaitanya S.					
; APPLICANT: Lodes, Michael A.					
; APPLICANT: Fanger, Gary					
; APPLICANT: Vedvick, Tom					
; APPLICANT: Carter, Darriack					
; APPLICANT: Retter, Marc					
Alignment Scores:					
Pred. No.:	2,63e-15	Length:	1012		
Score:	148.00	Matches:	30		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	100.00%	Indels:	0		
DB:	4	Gaps:	0		
US-09-684-215A-17 (1-30) x US-09-643-597-351 (1-1012)					





```
Alignment Scores:
Pred. No.: 4.5e-15 Length: 1557
Score: 148.00 Matches: 30
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-684-215A-17 (1-30) x US-09-620-412C-332 (1-1557)
QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProfile 20
DB 22 ACGGCGCGTCCGATAAATTCACAGCTGCCAGGTGGCAGGGATTCCGCCATTCCGATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
DB 82 GGCAGGCGATGCCGATCGGGCCAGATC 111

RESULT 14
US-09-598-419-332
: Sequence 332, Application US/09598419
: Patent No. 6565856
: GENERAL INFORMATION:
: APPLICANT: Skeiky, Yasir A.W.
: APPLICANT: Scholler, John
: TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
: TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
: FILE REFERENCE: 210121.469C6
: CURRENT APPLICATION NUMBER: US/09/598,419
: CURRENT FILING DATE: 2000-06-20
: NUMBER OF SEQ ID NOS: 357
: SOFTWARE: FastSeq for Windows Version 3.0/4.0
: SEQ ID NO 332
: LENGTH: 1557
: TYPE: DNA
: ORGANISM: Chlamydia trachomatis
US-09-598-419-332

Alignment Scores:
Pred. No.: 4.5e-15 Length: 1557
Score: 148.00 Matches: 30
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-684-215A-17 (1-30) x US-09-598-419-332 (1-1557)
QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProfile 20
DB 22 ACGGCGCGTCCGATAAATTCACAGCTGCCAGGTGGCAGGGATTCCGCCATTCCGATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
DB 82 GGCAGGCGATGCCGATCGGGCCAGATC 111

RESULT 15
US-09-556-877-188
: Sequence 188, Application US/09556877
: Patent No. 6432916
: GENERAL INFORMATION:
: APPLICANT: Probst, Peter
: APPLICANT: Bhatia, Ajay
: APPLICANT: Skeiky, Yasir
: APPLICANT: Fling, Steve
: APPLICANT: Maisonneuve, Jeff
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
: TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
: FILE REFERENCE: 210121.469C5
: CURRENT APPLICATION NUMBER: US/09/556,877
: CURRENT FILING DATE: 2000-04-19
: NUMBER OF SEQ ID NOS: 305

: SOFTWARE: FastSeq for Windows Version 3.0/4.0
: SEQ ID NO 188
: LENGTH: 1578
: TYPE: DNA
: ORGANISM: Chlamydia
US-09-556-877-188

Alignment Scores:
Pred. No.: 4.57e-15 Length: 1578
Score: 148.00 Matches: 30
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-684-215A-17 (1-30) x US-09-556-877-188 (1-1578)
QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProfile 20
DB 22 ACGGCGCGTCCGATAAATTCACAGCTGCCAGGTGGCAGGGATTCCGCCATTCCGATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
DB 82 GGCAGGCGATGCCGATCGGGCCAGATC 111

Search completed: September 5, 2003, 12:12:03
Job time : 20.0345 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 5, 2003, 08:27:24 ; Search time 518.276 Seconds  
(without alignments)  
1406.846 Million cell updates/sec

Title: US-09-684-215A-17  
Perfect score: 148  
Sequence: 1 TAASDNFQLSQGGQGAIPIGQAMAIAGQI 30

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlp  
-O/cgn2\_1/USPTO.spool.p/US09684215/runat\_05092003\_072202\_572/app\_query.fasta\_1.853  
-DB=EST\_QMWT-fastcap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=BLOSUM62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09684215.ecgn\_1.1.4575\_erunat\_05092003\_072202\_572 -NCPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_estl:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_eston:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
c 1	60	40.5	747	14	CB597827
c 2	57	38.5	238	10	BB077789
c 3	57	38.5	336	14	CB184974
4	57	38.5	343	14	CA930964
5	57	38.5	351	12	BI070561
6	57	38.5	366	12	BI055877
c 7	57	38.5	370	14	CA931030
c 8	57	38.5	370	14	CA931960
9	57	38.5	370	14	CA931997
10	57	38.5	370	14	CA932035
11	57	38.5	370	14	CA932108
12	57	38.5	370	14	CA932138
c 13	57	38.5	374	14	CA927770
c 14	57	38.5	374	14	CA928449
c 15	57	38.5	374	14	CA928493
c 16	57	38.5	374	14	CA928769
c 17	57	38.5	384	14	CA931116
c 18	57	38.5	385	14	CA931394
c 19	57	38.5	385	14	CA931501
20	57	38.5	387	14	CA931767
21	57	38.5	388	14	CA927636
c 22	57	38.5	388	14	CA928025
c 23	57	38.5	388	14	CA929028
c 24	57	38.5	389	14	CA927727
c 25	57	38.5	389	14	CA928142
26	57	38.5	390	14	CA927617
c 27	57	38.5	390	14	CA927744
c 28	57	38.5	390	14	CA927930
c 29	57	38.5	402	14	CA933902
c 30	57	38.5	403	13	BU824075
c 31	57	38.5	404	14	CA934515
c 32	57	38.5	408	14	CA930794
c 33	57	38.5	409	14	CA931302
c 34	57	38.5	409	14	CA931312
c 35	57	38.5	413	14	CA931446
c 36	57	38.5	413	14	CA931603
c 37	57	38.5	414	14	CA927769
c 38	57	38.5	414	14	CA928453
c 39	57	38.5	414	14	CA929036
40	57	38.5	415	13	BU819016
41	57	38.5	416	9	AI164854
42	57	38.5	417	14	CA928113
c 43	57	38.5	417	14	CA929039
c 44	57	38.5	427	14	CA931564
c 45	57	38.5	427	14	CA931687

# ALIGNMENTS

RESULT 1  
CB597827/c

LOCUS

DEFINITION

ACCSSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

1 (bases 1 to 747)

CB597827

AGENCY

IMAGE

CB597827

EST

GI:29515683

Mus musculus

(house mouse)

747 bp

linear

musculus CDNA clone

EST 03-APR-2003

CB597827

AGENCY

IMAGE

CB597827

EST

GI:29515683

Mus musculus

(house mouse)

Chordata; Vertebrata; Euteleostomi;

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 747)

Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata, Y., Shigemoto, S., Shinagawa, A., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomioka, N., Toya, T., Tsunoda, Y., Watahiki, A., Watanabe, S., Yamamura, T., Yamanaka, I., Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

RIKEN Mouse ESTs (Konno, H., et al.)

RIKEN Mouse ESTs (Konno, H., et al.)  
Unpublished  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216

Tel.: 81-43-503-3216  
 Fax: 81-43-503-3216  
 E-mail: genome-res@sc.riken.go.jp,  
 URL: <http://genome-gsc.riken.go.jp/>,  
 Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoaka, S., Sasaki,  
 N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Thermotranscription and thermostabilization and thermostabilization of full length  
 cDNA. *Proc. Natl. Acad. Sci. U.S.A.* 95 (2), 520-524 (1998)  
 Itoh, M., Katsunai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,  
 Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki,  
 Y. and Hayashizaki, Y. Automated filtration-based high-throughput plasmid preparation  
 system. *Genome Res.* 9 (5), 463-470 (1999)  
 Carninci, P. and Hayashizaki, Y.  
 High-efficiency full-length cDNA cloning. *Methods Enzymol.* 303,  
 19-44 (1999)  
 Please visit our web site (<http://genome.rtc.riken.go.jp/>) for  
 further details.

```

Location/Qualifiers
1. .338
  /organism="Mus musculus"
  /mol_type="mRNA"
  /strain="C57BL/6J"
  /db_xref="taxon:10090"
  /clone="9330145E04"
  /sex="male"
  /tissue_type="diencephalon"
  /dev_stage="adult"
  /lab_host="DH10B"
  /clone_lib="RIKEN full-length enriched, adult male

```

/Clone\_lib="RIKEN full-length enriched, adult male  
diencephalon"  
/note="Site1": Sali; Site2: BamHI; cDNA library was  
prepared and sequenced in Mouse Genome Encyclopedia  
Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in  
RIKEN. Division of Experimental Animal Research in Riken  
contributed to prepare mouse tissues. 1st strand cDNA was  
primed with a primer [5',  
GAGAGAGAGAGATCCACAGAGCTCTTTTTTTTTTNN 3'], cDNA was  
prepared by using trehalose thermo-activated reverse  
transcriptase and subsequently enriched for full-length by  
cap-trapper. cDNA went through one round of normalization  
to Rot = 10.0 and subtraction to Rot = 185.0. Second  
strand cDNA was prepared with the primer adapter of  
sequence [5', GAGAGAGAGATTCGAGTAAATTAATCCCCCCCCCCC  
3']. cDNA was cloned into the xhoI and BamHI sites.  
Vector: a modified pBluescript KS(+) after bulk excision  
from Lambda FLC I. Cloning sites, 5' end: Sali; 3' end:  
BamHI"

prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 185.0. Second strand cDNA was prepared with the primer adapter of sequence 5' GAGACAGAGATTCGAGTTAATTAAATTCCTCCCCCCCC 3'. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI<sup>a</sup>.

	47 a	77 c	48 g	66 t	
cores:	48.7				Length: 238
ilarity:	57.00				Matches: 9
ilarity:	75.00%				Conservative: 6
ilarity:	45.00%				Mismatches: 5
	38.51%				Indels: 0
	10				Gaps: 0

cores:	48.7	Length:	238
	57.00	Matches:	9
arity:	75.00%	Conservative:	6
ilarity:	45.00%	Mismatches:	5
	38.51%	Indels:	0
	10	Gaps:	0
			10



```

Email: riker@biochem.kth.se.
FEATURES
  source
    Location/Qualifiers
      1..351
        /organism="Populus tremula x Populus tremuloides"
        /mol_type="mRNA"
        /strain="T89"
        /db_xref="taxon:47664"
        /tissue_type="leaf"
        /clone_lib="Populus strain T89 leaves"
BASE COUNT 102 a 55 c 95 g 97 t 2 others
ORIGIN

Alignment Scores:
Pred. No.: 78.3 Length: 351
Score: 57.00 Matches: 10
Percent Similarity: 66.67% Conservative: 6
Best Local Similarity: 41.67% Mismatches: 8
Query Match: 38.51% Indels: 0
DB: 12 Gaps: 0

US-09-684-215A-17 (1-30) x BI070561 (1-351)
QY 1 ThrAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 32 ACATCTGGAGCAGAGGATTCGAAGTGTTCACAGGAGGATCAGGTTCCGCA 91
QY 21 GlyGlnAlaMet 24
Db 92 GGAGAAATGATG 103

RESULT 6
BI055877
LOCUS CM3-GN0330-120201-735-e08 GN0330 Homo sapiens cDNA, mRNA sequence.
DEFINITION CM3-GN0330-120201-735-e08 GN0330 Homo sapiens cDNA, mRNA sequence.
ACCESSION BI055877
VERSION BI055877.1 GI:14463407
KEYWORDS EST.
SOURCE Homo sapiens
  ORGANISM
    Homo sapiens
      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
      1 (bases 1 to 366)
      Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
      Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
      Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
      Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
      M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
      Simpson,A.J.
      Shotgun sequencing of the human transcriptome with ORF expressed
      sequence tags
      Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
      20202663
      10737800
      Contact: Simpson A.J.G.
      Laboratory of Cancer Genetics
      Ludwig Institute for Cancer Research
      Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
      Brazil
      Tel: +55-11-2704922
      Fax: +55-11-2707001
      Email: asimpson@ludwig.org.br
      This sequence was derived from the FAPESP/LICR Human Cancer Genome
      Project. This entry can be seen in the following URL
      (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM3&t2=CM3-GN0330-
      120201-735-e08&t3=2001-02-12&t4=1)
      Seq primer: puc 18 forward
      High quality sequence stop: 366.
      Location/Qualifiers
        1..366
          /organism="Homo sapiens"
          /mol_type="mRNA"
          /db_xref="taxon:9606"
          /dev_stage="Adult"
FEATURES
  source
    Location/Qualifiers
      1..366
        /organism="Populus tremuloides"
        /mol_type="mRNA"
        /db_xref="taxon:3693"
        /clone_lib="Aspen apex cDNA Library"
        /note="organ: apex"
BASE COUNT 70 a 111 c 64 g 125 t
ORIGIN

Alignment Scores:
Pred. No.: 83.5 Length: 370
Score: 57.00 Matches: 10
Percent Similarity: 66.67% Conservative: 6
Best Local Similarity: 41.67% Mismatches: 8
Query Match: 38.51% Indels: 0
DB: 14 Gaps: 0

/clone_lib="GN0330"
/note="Organ: placenta_normal; Vector: puc18; Site_1: SmaI
; Site_2: SmaI; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent
application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the pUC 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."
BASE COUNT 82 a 91 c 143 g 50 t
ORIGIN

Alignment Scores:
Pred. No.: 82.4 Length: 366
Score: 57.00 Matches: 10
Percent Similarity: 71.43% Conservative: 5
Best Local Similarity: 47.62% Mismatches: 6
Query Match: 38.51% Indels: 0
DB: 12 Gaps: 0

US-09-684-215A-17 (1-30) x BI055877 (1-366)
QY 3 AlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIleGlyGln 22
Db 184 GCACCGGAGGAGGACAGCTCGATCGCCCTGGCAAGGATTCGTGCTGCTGGCGAG 243
QY 23 Ala 23
Db 244 GCC 246

RESULT 7
CA931030
LOCUS MTU2TA.P1.A09 Aspen apex cDNA Library Populus tremuloides cDNA,
DEFINITION MTU2TA.P1.A09 Aspen apex cDNA Library Populus tremuloides cDNA,
ACCESSION CA931030
VERSION CA931030.1 GI:27419510
KEYWORDS EST.
SOURCE Populus tremuloides (quaking aspen)
  ORGANISM
    Populus tremuloides
      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
      ; eurosids I; Malpighiales; Salicaceae; Populus.
      1 (bases 1 to 370)
      Ranjan,P., Kuo,Y.-Y., Harding,S.A., Jiang,H., Joshi,C.P. and Tsai
      C-J.
      Expressed sequence tags from Aspen
      Unpublished
      Contact: Tsai C-J
      Plant Biotech Research Center
      Michigan Technological University, School of Forest Resources &
      Environmental Science
      1400 Townsend Drive, Houghton, MI 49931-1295, USA
      Tel: 906 487 2914
      Fax: 906 487 2915
      Email: chtsa@mtu.edu.
      Location/Qualifiers
        1..370
          /organism="Populus tremuloides"
          /mol_type="mRNA"
          /db_xref="taxon:3693"
          /clone_lib="Aspen apex cDNA Library"
          /note="organ: apex"
BASE COUNT 70 a 111 c 64 g 125 t
ORIGIN

Alignment Scores:
Pred. No.: 83.5 Length: 370
Score: 57.00 Matches: 10
Percent Similarity: 66.67% Conservative: 6
Best Local Similarity: 41.67% Mismatches: 8
Query Match: 38.51% Indels: 0
DB: 14 Gaps: 0

```

SOURCE  
Populus tremuloides (quaking aspen)

Eukaryota; Viridiplantae; Scenedesmus; Eubryophyta; Charophyta; Magnoliophyta; Magnoliopsida; Eudicotyledons; Core eudicots; Rosids



; eusoids I; Malpighiales; Salicaceae; Populus.

REFERENCE 1 (bases 1 to 374)  
AUTHORS Ranjan,P., Kao,Y.-Y., Harding,S.A., Jiang,H., Joshi,C.P. and Tsai ,C.-J.

TITLE Expressed sequence tags from Aspen

JOURNAL Unpublished

COMMENT Contact: Tsai C-J  
Plant Biotech Research Center  
Michigan Technological University, School of Forest Resources &  
Environmental Science  
1400 Townsend Drive, Houghton, MI 49931-1295, USA  
Tel: 906 487 2914  
Fax: 906 487 2915  
Email: chtsal@mtu.edu.

FEATURES source

1. .374  
Location/Qualifiers  
/organism="Populus tremuloides"  
/mol\_type="mRNA"  
/db\_xref="taxon:3693"  
/clone\_lib="Aspen root cDNA Library"  
/note="Organ: root"

BASE COUNT 69 a 108 c 66 g 131 t  
ORIGIN

Alignment Scores:

Pred. No.: 84.6 Length: 374  
Score: 57.00 Matches: 10  
Percent Similarity: 66.67% Conservative: 6  
Best Local Similarity: 41.67% Mismatches: 8  
Query Match: 38.51% Indels: 0  
DB: 14 Gaps: 0

US-09-684-215A-17 (1-30) x CA927770 (1-374)

Qy 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProile 20

Db 146 ACATCTGGAGCAGGAGTTCGAAGTGTTCACAGGAGGATCAGGGTTCGCATTTCTCGCA 87

Qy 21 GlyGlnAlaMet 24

Db 86 GGAGAAATGATG 75

RESULT 14

CA928449/c

LOCUS MTU6TR.P2.G12 Aspen root cDNA Library Populus tremuloides cDNA, 374 bp mRNA linear EST 30-DEC-2002

DEFINITION mRNA sequence.

ACCESSION CA928449

VERSION CA928449.1 GI:27416928

KEYWORDS EST.

SOURCE Populus tremuloides (quaking aspen)

ORGANISM Populus tremuloides  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

REFERENCE 1 (bases 1 to 374)  
AUTHORS Ranjan,P., Kao,Y.-Y., Harding,S.A., Jiang,H., Joshi,C.P. and Tsai ,C.-J.

TITLE Expressed sequence tags from Aspen

JOURNAL Unpublished

COMMENT Contact: Tsai C-J  
Plant Biotech Research Center  
Michigan Technological University, School of Forest Resources &  
Environmental Science  
1400 Townsend Drive, Houghton, MI 49931-1295, USA  
Tel: 906 487 2914  
Fax: 906 487 2915  
Email: chtsal@mtu.edu.

FEATURES source

1. .374  
Location/Qualifiers  
/organism="Populus tremuloides"  
/mol\_type="mRNA"  
/db\_xref="taxon:3693"

/clone\_lib="Aspen root cDNA Library"  
/note="Organ: root"

BASE COUNT 71 a 110 c 64 g 129 t  
ORIGIN

Alignment Scores:

Pred. No.: 84.6 Length: 374  
Score: 57.00 Matches: 10  
Percent Similarity: 66.67% Conservative: 6  
Best Local Similarity: 41.67% Mismatches: 8  
Query Match: 38.51% Indels: 0  
DB: 14 Gaps: 0

US-09-684-215A-17 (1-30) x CA928449 (1-374)

Qy 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProile 20

Db 146 ACATCTGGAGCAGGAGTTCGAAGTGTTCACAGGAGGATCAGGGTTCGCATTTCTCGCA 87

Qy 21 GlyGlnAlaMet 24

Db 86 GGGGAATGATG 75

RESULT 15

CA928493/c

LOCUS MTU6TR.P3.C12 Aspen root cDNA Library Populus tremuloides cDNA, 374 bp mRNA linear EST 30-DEC-2002

DEFINITION mRNA sequence.

ACCESSION CA928493

VERSION CA928493.1 GI:27416972

KEYWORDS EST.

SOURCE Populus tremuloides (quaking aspen)

ORGANISM Populus tremuloides  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

REFERENCE 1 (bases 1 to 374)  
AUTHORS Ranjan,P., Kao,Y.-Y., Harding,S.A., Jiang,H., Joshi,C.P. and Tsai ,C.-J.

TITLE Expressed sequence tags from Aspen

JOURNAL Unpublished

COMMENT Contact: Tsai C-J

Plant Biotech Research Center

Michigan Technological University, School of Forest Resources &

Environmental Science

1400 Townsend Drive, Houghton, MI 49931-1295, USA

Tel: 906 487 2914

Fax: 906 487 2915

Email: chtsal@mtu.edu.

FEATURES source

1. .374  
Location/Qualifiers

/organism="Populus tremuloides"

/mol\_type="mRNA"

/db\_xref="taxon:3693"

/clone\_lib="Aspen root cDNA Library"

/note="Organ: root"

BASE COUNT 70 a 110 c 65 g 129 t

ORIGIN

Alignment Scores:

Pred. No.: 84.6 Length: 374  
Score: 57.00 Matches: 10  
Percent Similarity: 66.67% Conservative: 6  
Best Local Similarity: 41.67% Mismatches: 8  
Query Match: 38.51% Indels: 0  
DB: 14 Gaps: 0

US-09-684-215A-17 (1-30) x CA928493 (1-374)

Qy 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProile 20

Db 146 ACATCTGGAGCAGGAGTTCGAAGTGTTCACAGGAGGATCAGGGTTCGCATTTCTCGCA 87

Qy 21 GlyGlnAlaMet 24  
   | | | : : :  
Db 86 GGAGAAATGATG 75

Search completed: September 5, 2003, 11:54:36  
Job time : 520.276 secs